



UNIVERSITÀ
DEGLI STUDI
DI BRESCIA

DIGITAL TRANSFORMATION

Evaluating the impact of ontology integration on RAG performance across two domains

Autors:

Filippini Elena

Badge n. 727828

Filippini Toninato Dario

Badge n. 727825

Accademic Year 2023/2024

Executive Summary

The paper explores the integration of an ontology into a Retrieval-Augmented Generation (RAG) system, aiming to enhance its performance in the domains of salmon and cybersecurity. By comparing a naive RAG with an ontology-enhanced version (RAGOnto), the study assesses whether the latter provides more accurate and relevant responses.

Key findings include:

- Objective: to evaluate if RAGOnto outperforms a naive RAG in domain-specific tasks.
- Methodology: testing was conducted in the salmon and cybersecurity domains using predefined queries, without training new Large Language Models.
- Ontology integration: ontologies were incorporated to refine query results by adding related entities to the questions.
- Results: RAGOnto showed improved performance, especially in the salmon domain, where the ontology was more aligned with the content. The cybersecurity results were less pronounced, highlighting the importance of ontology quality.
- Conclusion: ontology integration can enhance RAG performance, but its success depends on domain relevance and the structure of the ontology used. Further automation in this process is recommended to improve the actual results.

The study concludes that while ontology integration has potential, its effectiveness varies across domains and requires careful implementation.



Content from this work © 2024 by Filippini Toninato Dario, Filippini Elena is licensed under CC BY 4.0. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>.

Contents

Executive Summary	1
Introduction	4
1 Background	5
2 Methodology	7
2.1 Parameters setting	7
2.2 Ontology integration	8
3 Results	10
3.1 Salmon domain	10
3.1.1 Salmon RAG	10
3.1.2 Salmon RAGOnto	11
3.2 Cybersecurity domain	14
3.2.1 Cybersecurity RAG	14
3.2.2 Cybersecurity RAGOnto	16
4 Discussion	19
4.1 Cross-domain comparing	19
4.2 Independent domain observations	19
4.3 Limits of our approach	20
5 Conclusions	21
Acknowledgements	21
References	22
Appendices	27

A Legenda of metrics' queries	27
B Salmon RAG queries	27
C Salmon RAGOnto queries	47
D Cybersecurity RAG queries	70
E Cybersecurity RAGOnto queries	89

Introduction

The Large Language Models (LLMs) nowadays are a very powerful general-purpose technology which allow to solve many common tasks. Focusing on the question-answer functionality, an LLM, such as the most famous ChatGPT [1], follows this process: it reads the question made by the user, tries to understand its semantics, retrieves the most relevant information from the available data and, in the end, it generates what it considers the best answer. Because of their functioning, the success of LLMs is probabilistic. To increase such probability, it has been introduced Retrieval Augmented Generation (RAG) systems that usually rely on an already existing LLM but are specialized only in one domain. To make it possible, a lot of new documents about a single domain are made available to the RAG.

The documents used by RAG are in a textual format: one of the worst formats for a machine. So, to further improve the probability of a right answer, a RAG could also utilize well-formed documents beyond the textual ones. One possible structured data that could be exploited is ontologies.

To better understand the above concepts, let us consider an example .

Question: ‘Which is the salmon’s most common parasite?’

LLM answer: ‘Sea lice’ and list its most important characteristics.

RAG answer: ‘Sea lice’ and describe it in details, talking for example about its morphology and feeding habits.

RAG with ontology provides the same answer as RAG, along with additional information about the second most common salmon parasite: ‘Gyrodactylus salaris’.

1 Background

To appreciate the work presented in this paper, our audience should have a foundational understanding of:

- Natural Language Processing (NLP): the branch of artificial intelligence focused on the interaction between computers and humans through natural language.
- Retrieval-Based Models: systems designed to fetch relevant information from a corpus in response to a query.
- Generative Models: models that generate new data samples, particularly text, by learning from a training set of existing examples.
- Ontologies: structured representations of knowledge within a domain, defining concepts and their interrelationships.
- Domain-Specific Knowledge: understanding the basic terminologies of specialized fields such as salmon and cybersecurity.

Retrieval-Augmented Generation (RAG) is an advanced method in NLP that combines the strengths of retrieval-based models and generative models. In essence, RAG systems first retrieve relevant information from a large corpus and, then, use this information to generate a contextually appropriate answer [2]. However, to further improve the accuracy of retrieved information, RAG could be enhanced with ontologies. Naive RAG systems, while effective in many general scenarios, often fall short in specialized fields where understanding the domain-specific terminologies is crucial.

Ontologies provide a formal representation of knowledge within a specific domain, including the concepts and relationships among them. So the information is hierarchically structured and all kinds of relations and properties are well-defined, unlike textual documents.

By incorporating an ontology, the retrieval component of a RAG system can leverage the structured knowledge to fetch more contextually appropriate information. This, in turn, enhances the generation phase, leading to outputs that are factually aligned with the domain-specific requirements [3]. These findings are the main reason why our audience should read this paper, especially those who

need to develop customized RAGs for their company and want to maximize their effectiveness. We built our RAG application in Python using the Langchain framework. Following the examples in the tutorials section of the documentation [4], we created the first raw version of the system, which was then improved during the experimentation phase. After the documentation loading, it's necessary to define the main components of our RAG: the pipeline, the prompt template, and the retriever. The pipeline declares as LLM the 'Mistral-7B-v0.1' model [5] and as task the text generation. The prompt template is an instruction about how to answer the question. The role of the retriever is to obtain a specific number of chunks from the vector database after the original documents have been chunked and embedded. Our chunks contain one or more sentences, depending on their chosen size. Then, these chunks are embedded and loaded into the FAISS index (Facebook AI Similarity Search) [6]: the vector database integrated into the systems. In the end, the retriever is defined to connect the query to FAISS index and it is used with the pipeline and the prompt template to create the actual RAG. It's invoked by passing the query as a parameter to generate the answer based on the individuated context. It searches several document pieces chosen by us, giving the priority to the chunks that contain the most relative information to answer the question according to the cosine similarity criterion between the query and document embeddings.

2 Methodology

The main purpose of this work is to experimentally verify if a RAG with an ontology, called RAGOnto, can outperform a naive RAG or if their performance is fairly comparable. The ontologies utilize the Web Ontology Language (OWL) 2 EL designed by W3C. To achieve this goal, two specific oriented domains have been chosen: salmon and cybersecurity. We chose the cybersecurity domain because we believe it could benefit particularly from the use of ontologies [7] [8], as this domain is made of unstructured content, such as reports, ‘Capture The Flag’ write-ups and standards. The main reason behind the choice of the salmon domain is the ontology we found. In fact, the ontology [9] has an appropriate number of classes, about 1100, and its structure is not either too complex or too simple for our purposes. Then, a sample of fourteen queries for each domain has been selected to be submitted to both RAG and RAGOnto. The answers generated are compared and used as an evaluation yardstick through the metrics made available by ragas library [10].

Therefore, domains different from salmon and cybersecurity, are not considered. In addition to that, in our study RAG and RAGOnto don’t generalize on unexpected questions: they answer in the best possible way only to the selected queries. In the end, it isn’t our goal to train a new LLM or to create a new embedding method: we use the existing libraries in our systems.

2.1 Parameters setting

Once defined the RAG architecture, we tested its working on the first queries. During this testing period, we identified the most significant parameters: the size of a chunk (`chunk_size`), expressed as the maximum number of characters, and the chunk’s number (`k`) obtained by the retriever. We found that these parameters must be carefully adjusted for each query, as they are crucial to finding the correct context. The idea of ‘`chunk_size`’ is that the more the question is broad but specific to a document section, the greater its value should be. So, the minimum values are around 200-500 if the question refers to a few short sentences in the documents, and 4000-5000 if the query requires one or two pages of information. ‘`k`’ can be assigned a value between 1 and 3. We noticed that if ‘`k`’ is greater than 3 the system struggles to integrate the heterogeneous retrieved information, going out of context.

Finally, the way the query is formulated is also important. For example, asking ‘What are the Pacific salmon species?’ and ‘List all *Oncorhynchus* species.’ isn’t the same. Although the semantics of the questions is identical, the syntax and lexicon are different, so the retrieved context could be different. In this example, the meaningful changes are the starting (‘What’ vs ‘List’) and using ‘*oncorhynchus*’ as synonymous of ‘pacific salmon’. Therefore sometimes we had to change the syntax and lexicon of a query several times to obtain an acceptable answer.

2.2 Ontology integration

Now let’s see how the ontology has been integrated into the system. We searched for a similar work on the web to be taken as a reference, however, the only really useful thing that we found was about converting a textual document into a knowledge graph, that it’s quite similar to an ontology [11] [12]. This is the reversed process of our goal. So, the following described procedure is our original idea.

We identified one or more keywords of a query, based on its semantic, use them to get all the related information in the ontology and modify the original question adding this information. Keywords are entities within the ontology, as is the retrieved information. The point is to start from a key entity, obtain all the linked entities, and modify the current question by suggesting to the RAG to use them during its context-searching phase. The process to get the list of the related entities to a given one is the following: first, the key entity is included, then recursively all its subclasses and superclasses are added. The siblings of the key entity and those of its superclasses are excluded to reduce noise. However, ontologies also have properties, so for each considered class, all their properties are taken into account to get new entities linked to the key one. This example clarifies the process:

```

c1 Salmonid type
  c2 Pacific Trout
    c3 Cutthroat Trout
    c3 Steelhead Trout
      p hasSameDietOf Pink Salmon
c2 Pacific Salmon
  c3 Pink Salmon
    p hasSameDietOf Steelhead Trout

```

c3 Dog Salmon

Where c is the class, the number next to it indicates the level and p stands for the property. With ‘Pacific Salmon’ as the keyword, the entities added to the query are {Pacific Salmon, Pink Salmon, Steelhead Trout, Dog Salmon, Salmonid type}, exactly in this order.

This is the working of our RAGOnto. Using the above example, on a query such as ‘What is the diet of Pink Salmon?’, the best possible answer for a RAG is to correctly describe the diet, while it’s reasonable to think that RAGOnto can expand the answer by adding the detail that Steelhead trout has the same diet. The question becomes the following:

‘What is the diet of Pink Salmon? These concepts might be useful to answer the question:

Pink Salmon, Steelhead Trout, Pacific Salmon, Salmonid type’.

We would like to clarify that this method doesn’t filter the considered properties: they are all taken into account. Let’s say that instead ‘hasSameDietOf’, ‘Pink Salmon’ has the properties ‘hasSameColorOf’ referred to ‘Cutthroat Trout’ and ‘hasDifferentDietFrom’ still refers to ‘Steelhead Trout’. In the previous query, both ‘Cutthroat Trout’ and ‘Steelhead Trout’ are added as key entities: the RAG, based on the question semantic, should ignore the useless class ‘Cutthroat Trout’ and focus on ‘Steelhead Trout’.

3 Results

See [Appendix A] to understand the metrics shown in the table’s header.

All the metrics values are between 0 and 1, where 0 is the worst case and 1 is the best one.

The most relevant metrics are context recall, faithfulness, and answer similarity. The first metric demonstrates the robustness of the RAG, while the other metrics ensure its reliability.

3.1 Salmon domain

3.1.1 Salmon RAG

QUERY	CP	F	AR	CR	CER	AS	SS
1	1.0	1.0	0.907515	1.0	0.692308	0.956241	0.512712
2	1.0	0.0	0.843462	0.0	0.333333	0.898259	0.550121
3	1.0	0.333333	0.863908	1.0	0.428571	0.938321	0.714103
4	1.0	0.307692	0.917987	1.0	0.111111	0.936895	0.029466
5	1.0	0.571429	0.977609	1.0	0.083333	0.976953	0.541165
6	1.0	0.222222	0.924439	0.428571	0.111111	0.939084	0.496454
7	1.0	1.0	0.960842	0.666667	0.666667	0.955333	0.785882
8	1.0	0.333333	1.0	1.0	0.900000	0.942211	0.629323
9	1.0	0.583333	0.917076	0.750000	0.250000	0.964756	0.415197
10	1.0	0.920000	0.992133	1.0	0.521739	0.958139	0.540238
11	1.0	0.727273	0.918844	1.0	0.636364	0.978910	0.417718
12	1.0	0.538462	0.970617	0.875000	0.785714	0.933236	0.668773
13	1.0	0.333333	0.907109	0.0	0.0	0.818108	0.505682
14	0.0	0.285714	1.0	0.0	0.0	0.891360	0.401337
Mean	0.928571	0.511152	0.935824	0.694303	0.358589	0.934843	0.510660
Std	0.267261	0.308000	0.049152	0.412073	0.303859	0.042042	0.179917

Table 1: Metrics values for salmon domain’s RAG queries

From the yellow row of table 1, it is clear that context entity recall has the worst values. This means that the concepts expressed in the retrieved context and the ones in the ground truth are quite different. The possible explanations for this behavior are that the system gets a wrong context or the context contains too much or too little information. RAG parameters ‘chunk_size’ and ‘k’ are carefully set to minimize the first case. Also, faithfulness has quite low scores, meaning that the number of claims in the generated answer that can be inferred from the given context is far

less than the total number of claims in the generated answer. This happens because, to add details to the answer for some queries, the system uses data from the LLM’s extra context. For example, let’s analyze query [4]. The CER is only 0.111 because all the entities about ‘Traditional fisheries’ are not in the ground truth since they aren’t necessary to answer and instead the context lacks information about Alaska. Faithfulness is also low, 0.307, mainly because in the answer there is extra information about Alaska regulations that isn’t present in the context.

SS mean is 0.51, however, we would like to point out that summarizing the context is a secondary task for this type of RAG. Comparing the ground truth with the given answer, we realized that for answer similarity a value below the threshold of 0.9 describes an incomplete or wrong answer; the same is true for the answer relevancy metric. Let us examine queries [13] and [2] as examples. Despite the query 13 answer being wrong, AS has still a value of 0.818. AR is calculated taking into account the question, the context, and the answer: for query 2 its value is 0.843, the minimum among all queries. The answer is right, but incomplete and it lacks details.

3.1.2 Salmon RAGOnto

QUERY	CP	F	AR	CR	CER	AS	SS
1	1.0	1.0	0.907515	1.0	0.692308	0.956241	0.512712
2	1.0	0.0	0.843462	0.0	0.333333	0.898259	0.550121
3	1.0	0.571429	0.938846	0.0	0.142857	0.934211	0.537153
4	1.0	0.875000	0.934789	1.0	0.789474	0.963741	0.444444
5	1.0	0.666667	0.877814	1.0	0.166667	0.953670	0.425632
6	1.0	0.428571	0.924439	0.375000	0.111111	0.984627	0.639275
7	1.0	0.971429	0.961373	1.0	0.0	0.957691	0.499920
8	1.0	1.0	1.0	1.0	1.0	0.940686	0.911201
9	1.0	0.833333	0.917076	0.750000	0.153846	0.961277	0.499554
10	1.0	0.920000	0.992133	1.0	0.521739	0.958139	0.540238
11	1.0	0.727273	0.918844	1.0	0.636364	0.978910	0.417718
12	1.0	0.538462	0.970617	0.875000	0.785714	0.933236	0.668773
13	1.0	1.0	0.998634	1.0	0.800000	0.923555	0.738628
14	1.0	0.333333	0.967355	0.666667	0.400000	0.930889	0.841868
Mean	1.0	0.704678	0.939493	0.761905	0.466672	0.948224	0.587660
Std	0.0	0.301632	0.045893	0.370676	0.321466	0.022943	0.152987

Table 2: Metrics values for salmon domain’s RAGOnto queries

Observing the mean of the metrics in table 2, it can be seen how all these values have been improved compared to table 1: RAGOnto works better than RAG on the salmon domain. In

particular, the average faithfulness score increases by 20%, demonstrating that the use of the retrieved entities in the question increases the likelihood of generating an answer based only on the context rather than on extra information. The worst metric for RAGOnto remains the CER, although it's improved by 10%: the ontology entities slightly help in retrieving a context closer to the ground truth.

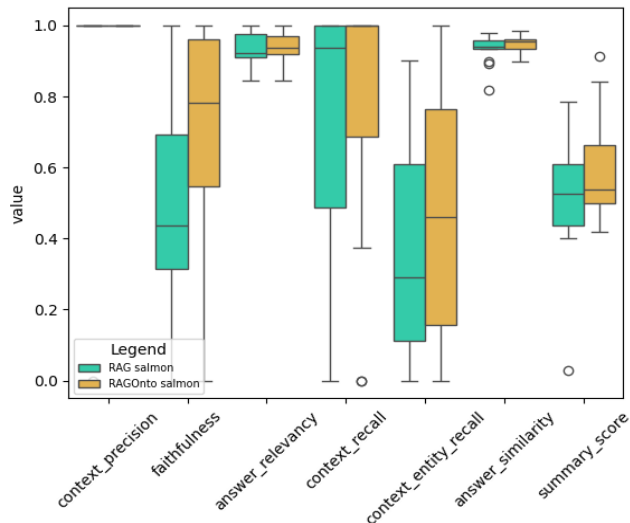


Figure 1: Comparing salmon RAG and RAGOnto’s values on each metric

As shown in table 2 and in figure 1, metrics with the best values’ distribution are CP, AR, and AS, while metrics with the worst ones are F, CR and CER. The reason is that a lot of their values are near to 1 or near to 0, lacking consistency. Furthermore, the behavior of standard deviation is the same for both RAG and RAGOnto, despite the last one increase a bit the metrics’ means and improve the standard deviation of CR.

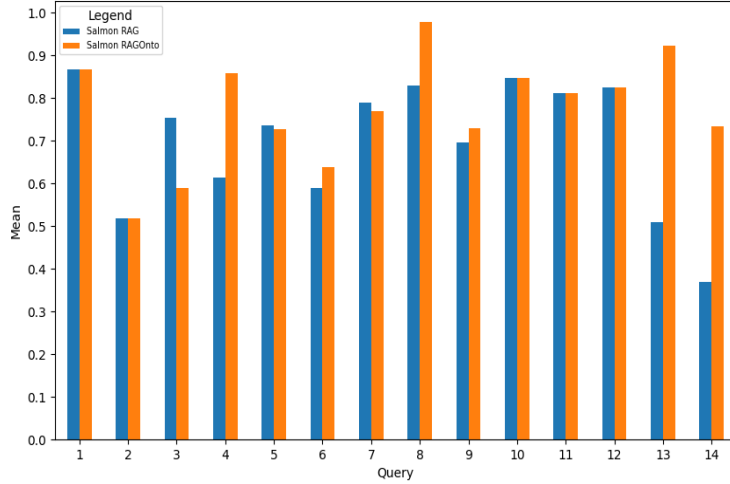


Figure 2: Comparing salmon RAG and RAGOnto’s means of metrics on each query

As it can be seen in figure 2, six queries improve their results with RAGOnto, five stay the same and three are worst, even if the only relevant one is query [3]. This happens because of CR, meaning that in this specific case entities retrieved aren’t helpful. However, overall the figure 2 confirms that RAGOnto performs better than RAG. In particular, the relevant improvements are in query [4], [13] and [14].

The RAGOnto query [4] retrieves a different context from the RAG query [4]; it contains both the definition of subsistence fishery and its regulations in Alaska. In this way, AR goes from 0.917 to 0.934 and CER from 0.111 to 0.789. Thanks to the right context, the RAGOnto answer gives the right information about Alaska prohibitions, unlike the RAG answer.

After analyzing query 13, it is clear how poorly works RAG [13] compared to RAGOnto [13]. First of all, RAG retrieves a wrong context, deceived by the similar semantics of ‘healthy product’ and ‘sustainable properties’. However, there is another factor that has tricked the RAG: the use of ‘salmo salar’ instead of ‘atlantic salmon’ in the question. RAGOnto can recognize ‘salmo salar’ as synonymous with ‘Atlantic salmon,’ whereas the naive RAG cannot. So RAGOnto can identify the right context and generate the appropriate answer.

Similar reasoning can be done for the query [14]. Although the answer generated by RAGOnto is way better than the one of the RAG, it isn’t perfect: fish stock code is considered a fish stock type, but it isn’t.

Even though RAGOnto behaves generally better than RAG, this doesn't mean that it is error-free. Its main quality, the use of the entities in the question, could become its weakness, especially in the queries that have been added many entities: some of these could be misleading, guiding RAGOnto to an inaccurate context. This is why in the answer to question [14] there is that mistake. The query [5], one with the greater number of entities, thirty-four, is the most striking example. What is asked of the system is information about internal tags, however part of the context refers to the external ones. In fact, some of the entities are about external tags, so actually RAGOnto does a good job in providing this kind of information. Despite the inaccurate context, the semantics of the question with the right entities, like 'pit tag', help in the identification of a quite correct answer. As it is observable in figure 2, this query achieve slightly better results with the naive RAG.

3.2 Cybersecurity domain

3.2.1 Cybersecurity RAG

QUERY	CP	F	AR	CR	CER	AS	SS
1	1.0	0.5	0.931145	1.0	0.249999	0.955379	0.647196
2	1.0	0.228571	0.940226	0.500000	0.166666	0.930849	0.0967993
3	1.0	0.0	0.0	0.666666	0.071428	0.832501	0.027257
4	1.0	0.916666	0.949954	0.333333	0.117647	0.963552	0.284836
5	1.0	0.181818	0.942531	0.333333	0.400000	0.952234	0.434135
6	1.0	0.666666	0.954977	1.0	0.250000	0.944036	0.579782
7	1.0	0.571429	0.860158	0.750000	0.142857	0.954277	0.669046
8	1.0	0.384615	0.935247	0.750000	0.875000	0.911144	0.019500
9	1.0	1.0	0.991193	0.500000	0.230769	0.972349	0.608546
10	1.0	0.666667	0.969668	1.0	0.076923	0.897330	0.520845
11	1.0	0.666667	0.983176	1.0	0.375000	0.924146	0.541667
12	1.0	0.333333	0.991535	1.0	0.111111	0.957907	0.680891
13	1.0	0.500000	0.915044	0.0	0.0	0.967812	0.108601
14	1.0	0.0	0.913923	0.666667	0.055556	0.943414	0.432035
Mean	1.0	0.472602	0.877056	0.678571	0.223068	0.936209	0.403653
Std	0.0	0.306034	0.254840	0.314970	0.221246	0.036854	0.247869

Table 3: Metrics values for cybersecurity domain's RAG queries

It can be observed from the above table's second-last line that context entity recall has again the lowest score among all the metrics, while answer similarity and context precision have the highest values.

Let us take into account queries [6] and [4] that ask the same thing, but the latter has an additional question about the definition of RAT. Because of this additional question, the RAG retrieves two different contexts for the queries, leading a two distinct answers. Faithfulness of query [4] is way better then the one of query [6] since query [6] context contains too much information. However context recall is equal to 1 for the query [6], while is only 0.333 for query [4]: in fact the RAT' definition there isn't in the context, while is present in the ground truth. So, adding an extra question in a query requires a minimum of caution since it can influence the answers to the previous questions because the retrieved context from the system can change.

Query [12] has the highest AR and AS values, while F and CER are quite low. Its answer gives more details than the ground truth. Even though important concepts are mentioned, they are not elaborated on as thoroughly as in the ground truth.

Query [3] achieves the worst results: faithfulness, answer_relevancy, and context_entity_recall are all close to zero. The fact that it achieves the worst result isn't surprising, because the RAG begins by saying it doesn't know the answer, then proceeds to give general information about how to protect yourself from malicious software and cyber threats.

3.2.2 Cybersecurity RAGOnto

QUERY	CP	F	AR	CR	CER	AS	SS
1	1.0	0.333333	0.944766	1.0	0.249999	0.920822	0.425233
2	1.0	0.0	0.914872	1.0	0.124999	0.678549	0.500000
3	1.0	0.0	0.972375	1.0	0.199999	0.955618	0.397904
4	1.0	1.0	0.891759	0.800000	0.083333	0.921999	0.479586
5	1.0	0.647059	0.958455	0.666667	0.600000	0.960561	0.354268
6	1.0	1.0	0.966252	1.0	0.375000	0.956029	0.761321
7	1.0	0.571429	0.860158	0.750000	0.142857	0.954277	0.669046
8	1.0	0.384615	0.935247	0.750000	0.875000	0.911144	0.019500
9	1.0	1.0	0.991193	0.500000	0.230769	0.972349	0.608546
10	1.0	0.588235	0.944719	1.0	0.076923	0.911857	0.371425
11	1.0	0.666667	0.777533	1.0	0.250000	0.953617	0.605442
12	1.0	0.333333	0.991535	1.0	0.111111	0.957907	0.680891
13	1.0	0.4	0.923855	0.666667	0.750000	0.977412	0.337264
14	1.0	0.750000	0.913923	0.500000	0.277778	0.972484	0.576017
Mean	1.0	0.590359	0.927617	0.830952	0.310555	0.928902	0.484746
Std	0.0	0.330634	0.056803	0.193578	0.253466	0.075490	0.189447

Table 4: Metrics values for cybersecurity domain’s RAGOnto queries

We chose two different ontologies for the cybersecurity domain because neither MALOnto nor UCO alone provides enough information relative to our documents. MALOnto offers insights into malware and related concepts; UCO stands for ‘Unified Cybersecurity Ontology’. Given that MALOnto is somewhat more extensive than UCO, we used MALOnto for eight queries and UCO for six.

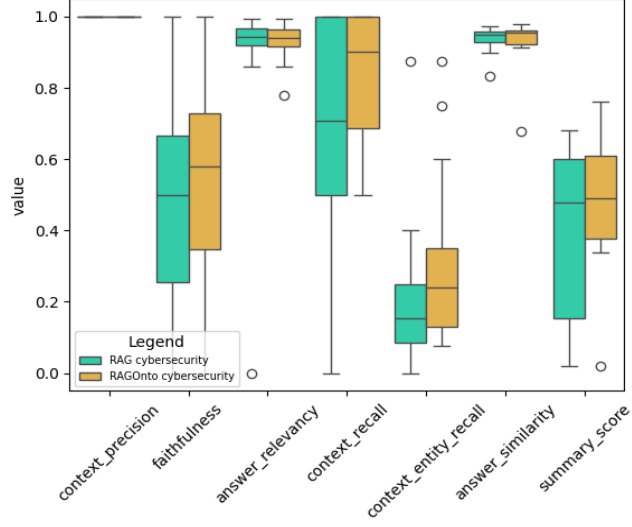


Figure 3: Comparing cybersecurity RAG and RAGOnto's values on each metric

As shown in table 4 and in figure 3, the standard deviation behavior is quite similar to the one of the salmon domain, but with two remarkable differences. The first one is that CER distribution is better, even if its mean is lower; the second one is that both standard deviation and means of CR and SS are improved with RAGOnto, proving that the entities of MAL and UCO ontologies at least help in finding more precise contexts.

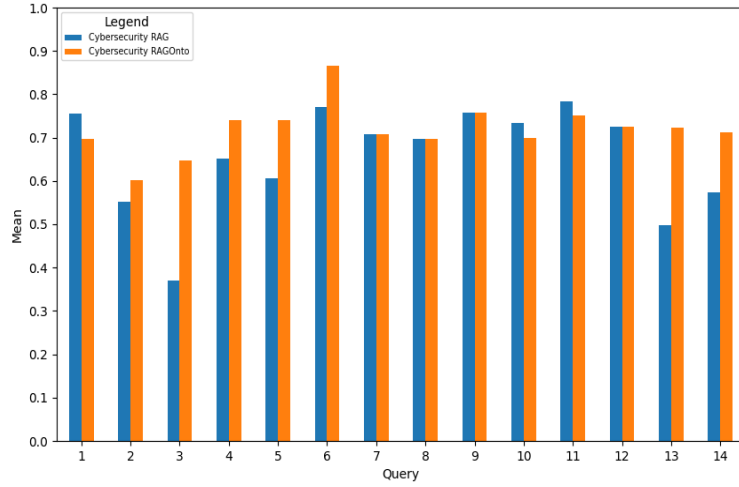


Figure 4: Comparing cybersecurity RAG and RAGOnto's means of metrics on each query

As it can be seen in figure 4, on average the queries have a higher score with RAGOnto. Even if

some queries' means worsened slightly, the improvements on other queries are more remarkable. In particular, queries with low mean values benefit more from the ontology compared to queries with already good results. The most meaningful increases are in queries [3], [13] and [14]. We would like to point out that some queries, such as [7] or [12], both have the same performance in RAGOnto and RAG, meaning that the entities taken into account don't add useful information. Finally, from comparing figure 2 and 4, can be noticed that cybersecurity queries' means are lower than the ones of salmon: RAG and RAGOnto performances are better in this domain.

Let us use query [13] as example. Despite the RAG answer being relevant, its metrics related to the context are particularly low. The context contains an appendix section, but it doesn't recall the ground truth concepts. Conversely, even if it's not the best possible one, the context of RAGOnto contains some information useful to answer. The reason the two systems give the same answer is that RAG retrieves data from an additional context rather than the given context. The same happens for the query [14]. This behavior must be avoided because the answer generated in this way is unpredictable.

Overall, the queries obtained with UCO ontology achieve better results in general, even though query [6] with MALOnto has the absolute highest value. This may be due to the UCO ontology being more detailed and aligned with the provided documents

4 Discussion

4.1 Cross-domain comparing

Comparing table 2 with table 4, it can be observed that almost all the metrics’ averages of salmon RAGOnto are greater than the ones of cybersecurity RAGOnto. This indicates that the quality improvement achieved by salmon RAGOnto is more substantial: in fact, it corrects incorrect or incomplete responses more frequently. The main reason for this is that the salmon ontology is more extended and with more entities related to salmon documents than the cybersecurity one. As a result, the faithfulness of salmon RAGOnto is 70%, compared to the 59% of cybersecurity RAGOnto, meaning that salmon ontology helps align the given answer with the context more effectively. However, we noticed that cybersecurity RAGOnto has the highest context recall, 83%, compared to 76% of salmon RAGOnto. Therefore cybersecurity ontologies are slightly more useful in identifying a context consistent with the ground truth.

4.2 Independent domain observations

All the observations made so far are specific to our domains; however, there are also domain-independent claims such as the best questions structure for a RAG and the type of queries where the ontology gives significant help. Typically, a RAG better understands questions that use words similar to those in the documents, and multiple simple questions are more effective than a single complex question. To obtain the definition of a concept, it’s better to ask for it explicitly in the query. Questions starting with ‘What are’ are very common and we realized that for such queries the system almost always tries to generate an answer, even if it’s wrong. To obtain the correct answer the rest of the questions must be logically formulated. Example of well formulated queries for a RAG are [1], [3], [11] about salmon domain and [1], [5], [7] about cybersecurity domain. Regarding ontology entities, we are confident that their effectiveness is high when they are used in classification questions or when seeking to find hidden connections between concepts. This effectiveness stems from the ontology’s hierarchical structure and the presence of properties, which enable the discovery of indirect relationships. Example are queries [7], [8], [14] about salmon domain.

Finally, it is important to address why the semantics of the ontology’s properties were not considered in RAGOnto. The reason is that no classes with disjoint properties has been added to the queries. Even if such classes exist in the three ontologies, they do not represent meaningful concepts. The commonly used disjoint property ‘DisjointWith’ in the ontologies is not employed. Furthermore, no instances of the ‘disconnected_from’ property, declared in salmon ontology [9], have been found. If, instead, the classes would use disjoint properties, the lack of semantic property filtering could result in both classes representing a specific concept and those representing the opposite concept being included in the same query, potentially causing confusion in RAGOnto.

4.3 Limits of our approach

The limits of this approach are the following.

- This way of integrating ontologies with RAGs is just one of the possible approaches and it may not be the optimal one.
- The method that retrieves the entities from the ontology must be adapted for each specific ontology.
- For each query, parameters ‘chunk_size’ and ‘k’ must be calibrated and this requires several executions.
- We identified the keywords by looking at the main concept in the question.
- It is clear that if the keywords in the question don’t match any entity in the ontology, then the use of such ontology does not provide any advantages. Conversely, if a keyword is connected to a large number of entities is likely that some of them are not relevant to the concepts in the question. So, the query risks going out of context since all the entities are used.
- It is advisable to introduce a semantic property filter that excludes disjoint properties if you apply this approach to an ontology containing such properties.

5 Conclusions

This analysis shows that the use of an ontology in a RAG without any regard, doesn't guarantee an improvement of the system and, even if there is, it can be only marginal. Indeed, achieving better results depends on several factors, such as the domain considered and how the concepts in the documents are systematically represented in the ontology through classes and their properties. We pointed out that RAGOnto works better in the salmon domain than in the cybersecurity one because salmon ontology is more complete and more linked to the concepts in the documents than the cybersecurity one. In conclusion, our study shows that, with the use of a well-structured ontology on a certain domain, RAG performance improves.

Our process to use the ontology in an RAG could be further improved. The RAGOnto could be automatized and generalized to all kinds of queries. To automatize it, first it should be developed an automatic method to extract the keywords to search the entities in the ontology; a possible way could be to retrieve the keywords from the main concepts in the question and/or in the ground truth. Then, should be handled the automatic settings of the 'chunk_size' and 'k' parameters, for example trying different values until an acceptable answer is given. However, we note that this solution could be expensive in terms of runtime and GPU memory. In the end, also a semantic filter about the retrieved ontology's entities could be added.

Acknowledgements

The advanced natural language processing capabilities of ChatGPT, developed by OpenAI, have been instrumental in facilitating the drafting, reviewing, and refining of the report. [1]

References

- [1] *ChatGPT*. URL: <https://chatgpt.com/> (visited on 08/30/2024).
- [2] Rick Merritt. *What Is Retrieval-Augmented Generation aka RAG?* NVIDIA Blog. Nov. 15, 2023. URL: <https://blogs.nvidia.com/blog/what-is-retrieval-augmented-generation/> (visited on 09/01/2024).
- [3] Plaban Nayak. *Implement RAG with Knowledge Graph and Llama-Index*. Medium. Apr. 30, 2024. URL: <https://medium.aiplanet.com/implement-rag-with-knowledge-graph-and-llama-index-6a3370e93cdd> (visited on 09/01/2024).
- [4] *Build a Retrieval Augmented Generation (RAG) App | LangChain*. URL: <https://python.langchain.com/v0.2/docs/tutorials/rag/> (visited on 08/30/2024).
- [5] *mistralai/Mistral-7B-v0.1 · Hugging Face*. URL: <https://huggingface.co/mistralai/Mistral-7B-v0.1> (visited on 08/30/2024).
- [6] *facebookresearch/faiss*. Aug. 30, 2024. URL: <https://github.com/facebookresearch/faiss> (visited on 08/30/2024).
- [7] *MALOnt - an ontology for Malware Threat Intelligence*. URL: <https://github.com/aiforsec/MALOnt/blob/main/MALOnt.owl> (visited on 08/30/2024).
- [8] *Unified Cyber Ontology*. URL: <https://github.com/ucoProject/UCO/blob/master/ontology/owl/owl.ttl> (visited on 08/30/2024).
- [9] *Salmon Ontology - Summary | NCBO BioPortal*. URL: <https://bioportal.bioontology.org/ontologies/SALMON/?p=summary> (visited on 08/28/2024).
- [10] *Introduction | Ragas*. URL: <https://docs.ragas.io/en/latest/index.html> (visited on 08/30/2024).
- [11] *Using a Knowledge Graph to Implement a RAG Application*. URL: <https://www.datacamp.com/tutorial/knowledge-graph-rag> (visited on 08/30/2024).
- [12] Rahul Nayak. *How to Convert Any Text Into a Graph of Concepts*. Medium. Nov. 17, 2023. URL: <https://towardsdatascience.com/how-to-convert-any-text-into-a-graph-of-concepts-110844f22a1a> (visited on 08/30/2024).

- [13] *Context Precision* / Ragas. URL: https://docs.ragas.io/en/latest/concepts/metrics/context_precision.html (visited on 08/30/2024).
- [14] *Faithfulness* / Ragas. URL: <https://docs.ragas.io/en/latest/concepts/metrics/faithfulness.html> (visited on 08/30/2024).
- [15] *Answer Relevance* / Ragas. URL: https://docs.ragas.io/en/latest/concepts/metrics/answer_relevance.html (visited on 08/30/2024).
- [16] *Context Recall* / Ragas. URL: https://docs.ragas.io/en/latest/concepts/metrics/context_recall.html (visited on 08/30/2024).
- [17] *Context entities recall* / Ragas. URL: https://docs.ragas.io/en/latest/concepts/metrics/context_entities_recall.html (visited on 08/30/2024).
- [18] *Answer semantic similarity* / Ragas. URL: https://docs.ragas.io/en/latest/concepts/metrics/semantic_similarity.html (visited on 08/30/2024).
- [19] *Summarization Score* / Ragas. URL: https://docs.ragas.io/en/latest/concepts/metrics/summarization_score.html (visited on 08/30/2024).
- [20] *Salmon*. In: *Wikipedia*. July 21, 2024. URL: <https://en.wikipedia.org/w/index.php?title=Salmon&oldid=1235917201> (visited on 08/28/2024).
- [21] *Resources*. Mowi Company Website. URL: <https://mowi.com/investors/resources/> (visited on 08/28/2024).
- [22] Adam Ridhwan Sukiman et al. "Overview of Machine Vision on Digital Imaging Approach for Automatic Tuna Length Measurement". In: *IOP Conference Series: Materials Science and Engineering* 551 (Aug. 2019). Publisher: IOP Publishing. URL: <https://dx.doi.org/10.1088/1757-899X/551/1/012076> (visited on 08/28/2024).
- [23] *Salmon Life Cycle* / Marine Institute. URL: <https://www.marine.ie/site-area/areas-activity/fisheries-ecosystems/salmon-life-cycle> (visited on 08/28/2024).
- [24] *UN Atlas of the Oceans: Subtopic*. URL: <https://www.oceansatlas.org/subtopic/en/c/1303/> (visited on 08/28/2024).

- [25] Fisheries {and} Oceans Canada Government of Canada. *Commercial salmon gear types / Pacific Region / Fisheries and Oceans Canada*. Last Modified: 2021-06-25. May 14, 2013. URL: <https://www.pac.dfo-mpo.gc.ca/fm-gp/salmon-saumon/comm-gear-engin-eng.html> (visited on 08/28/2024).
- [26] *Fishing gears*. URL: https://fish-commercial-names.ec.europa.eu/fish-names/fishing-gears_en (visited on 08/28/2024).
- [27] *Fish tags / ontario.ca*. URL: <http://www.ontario.ca/page/fish-tags> (visited on 08/28/2024).
- [28] *Getting Started*. URL: <https://www.oritag.org.za/GettingStarted> (visited on 08/28/2024).
- [29] RecordedFuture. *Chinese State-Sponsored RedJuliett Intensifies Taiwanese Cyber Espionage via Network Perimeter Exploitation*. URL: <https://www.recordedfuture.com/research/redjuliett-intensifies-taiwanese-cyber-espionage-via-network-perimeter> (visited on 08/28/2024).
- [30] *Confucius APT Android Spyware Linked to India-Pakistan Conflict / Threat Intel*. URL: <https://www.lookout.com/threat-intelligence/article/lookout-discovers-novel-confucius-apt-android-spyware-linked-to-india-pakistan-conflict> (visited on 08/28/2024).
- [31] *Nation-State Mobile Malware Targets Syrians with COVID-19 Lures / Threat Intel*. URL: <https://www.lookout.com/threat-intelligence/article/nation-state-mobile-malware-targets-syrians-with-covid-19-lures> (visited on 08/28/2024).
- [32] *Report: KL_report_syrian_malware*. URL: https://www.threatminer.org/report.php?q=KL_report_syrian_malware.pdf&y=2014 (visited on 08/28/2024).
- [33] *Russian Cyber Attack Campaigns and Actors*. GitHub. Rights note: ‘IronNet grants you a limited, revocable, non-exclusive, non-transferable license to access and use this Website for your personal use only. Neither this Website nor any of its content may be published, transmitted, copied, reproduced, resold or otherwise exploited, in whole or in part, for any commercial purpose without the express written consent of IronNet.’ URL: https://github.com/CyberMonitor/APT_CyberCriminal_Campagin_Collections/blob/master/2020/2020.

05.29_russian-cyber-attack-campaigns-and-actors/Russian%20Cyber%20Attack%20Campaigns%20and%20Actors%20-%20Threat%20Research.pdf (visited on 08/28/2024).

- [34] *Midyear Cybersecurity Report*. Cisco. Rights note: ‘Cisco routinely authors white papers, implementation guides, and technology overviews. Cisco does not generally license these materials for reproduction or distribution for commercial purposes (e.g., in for-profit books or compilations) or permit local hosting of these materials on third-party webpages. Instead, Cisco encourages and specifically authorizes parties to “deep link” to webpages on www.cisco.com, without additional permission from Cisco. Special exception for use in bona fide educational programs: Instructor duplication and distribution to students in bona fide educational programs of Cisco white papers, implementation guides, and technology overviews is expressly authorized, without additional permission from Cisco.’ 2017. URL: https://www.cisco.com/c/dam/m/digital/elq-cmcglobal/witb/1456403/Cisco_2017_Midyear_Cybersecurity_Report.pdf?elqTrackId=f6ccd8439e9945639096a9846044695a&elqaid=5897&elqat=2 (visited on 08/28/2024).
- [35] *Annual Cybersecurity Report*. Cisco. Rights note: ‘Cisco routinely authors white papers, implementation guides, and technology overviews. Cisco does not generally license these materials for reproduction or distribution for commercial purposes (e.g., in for-profit books or compilations) or permit local hosting of these materials on third-party webpages. Instead, Cisco encourages and specifically authorizes parties to “deep link” to webpages on www.cisco.com, without additional permission from Cisco. Special exception for use in bona fide educational programs: Instructor duplication and distribution to students in bona fide educational programs of Cisco white papers, implementation guides, and technology overviews is expressly authorized, without additional permission from Cisco.’ 2018. URL: <https://www.cisco.com/c/dam/m/digital/elq-cmcglobal/witb/acr2018/acr2018final.pdf?dtid=odicdc000016&ccid=cc000160&oid=anrsc005679&ecid=8196&elqTrackId=686210143d34494fa27ff73da9690a5b&elqaid=9452&elqat=2> (visited on 08/28/2024).
- [36] *Incident Response 2024 Report*. Palo Alto Networks. Rights note: ‘Palo Alto Networks grants you a limited, revocable, non-transferable license under its and other applicable copyrights to (1) download or use one copy of Site Content to a single computer, if applicable, solely

for your personal and non-commercial internal use.' URL: <https://www.paloaltonetworks.com/resources/research/unit-42-incident-response-report> (visited on 08/28/2024).

Appendices

A Legenda of metrics' queries

QUERY: the tuple question, ground truth, and answer. Each query can be seen in detail through the code link in the appendix.

CP (Context Precision): evaluates whether all of the ground-truth relevant items present in the contexts are ranked higher or not. [13]

F (Faithfulness): measures the consistency of the generated answer against the given context. [14]

AR (Answer Relevancy): focuses on assessing how pertinent the generated answer is to the given prompt. [15]

CR (Context Recall): measures the extent to which the retrieved context aligns with the ground truth. [16]

CER (Context Entity Recall): gives the measure of recall of the retrieved context, based on the number of entities present in both ground_truth and context relative to the number of entities present in the ground_truth alone. [17]

AS (Answer Similarity): measures the assessment of the semantic resemblance between the generated answer and the ground truth. [18]

SS (Summarization Score): gives a measure of how well the summary captures the important information from the context. [19]

B Salmon RAG queries

1. **Question:** Give a definition of seine net. What is the purpose of pair seines in fishing?

Context: Seines

A seine net is a very long net, with or without a bag in the centre, which is set either from the shore or from a boat for surrounding a certain area and is operated with two (long) ropes fixed to its ends (for hauling and herding the fish)Pair seines (SPR)

A seine net designed for pair fishing, where two boats operate the net in pair

Ground Truth: A seine net is a type of fishing net that hangs vertically in the water with its bottom edge held down by weights and its top edge buoyed by floats. It is a very long net,

sometimes featuring a bag in the center, which is deployed either from the shore or from a boat to surround a designated area. The net is operated using two ropes attached to its ends. Pair seines are designed for pair fishing, where two boats operate the net in pair.

Answer: A seine net is a very long net, with or without a bag in the centre, which is set either from the shore or from a boat for surrounding a certain area and is operated with two (long) ropes fixed to its ends (for hauling and herding the fish). Pair seines (SPR) are seine nets designed for pair fishing, where two boats operate the net in pair.

2. **Q:** What fishing net prevents fish from escaping using netting panels?

C: Surrounding nets and lift nets

Surrounding nets are large netting walls set for surrounding aggregated fish both from the sides and from underneath, thus preventing them from escaping by diving downwards

GT: Surrounding nets and lift nets prevent fish from escaping using netting panels.

A: Surrounding nets and lift nets

3. **Q:** What are trollers in commercial salmon fisheries?

C: Troller

The commercial troll fishery open area varies from year to year depending on the species available. For instance, if southern bound chinook salmon are the target species, then the west coast of Haida Gwaii is a location where northern trollers can harvest these chinook. On the north coast, the first commercial gill net fishery occurs May 31 in Area 8. Commercial net openings then occur on a weekly basis until mid October. Gill netters generally fish near coastal rivers and inlets, taking about 25% of the commercial catch.

Gill net vessel

Gill net

Trollers* employ hooks and lines which are suspended from large poles extending from the fishing vessel. Altering the type and arrangement of lures used on lines allows various species to be targeted. Trollers catch approximately 25% of the commercial harvest and finances *
Science and innovation

You are here:

1. Canada.ca 2. Fisheries and Oceans Canada 3. Pacific Region 4. Fisheries management 5. Pacific salmon 6. Commercial Pacific salmon fisheries

Commercial salmon gear types

Commercial salmon licences are issued for three gear types: seine, gill net and troll.

Seine nets* are set from fishing boats with the assistance of a small skiff. Nets are set in a circle around aggregations of fish

GT: Trollers in commercial salmon fisheries use hooks and lines suspended from large poles extending from the fishing vessel to catch pelagic fishes.

A: Trollers are a type of commercial salmon fishery that use hooks and lines to catch fish. They are typically used in areas where there are high concentrations of fish, such as near river mouths or inlets. Trollers can also be used in deeper waters, but they are not as effective at catching fish in those areas.

4. **Q:** What is subsistence fishery? Is there any prohibition in Alaska?

C: Subsistence fisheries

A fishery where the fish caught are shared and consumed directly by the families and kin of the fishers rather than being bought by intermediaries and sold at the next larger market. Pure subsistence fisheries are rare as part of the products are often sold or exchanged for other goods or services

Traditional fisheries

Fisheries established long ago, usually by specific communities that have developed customary patterns of rules and operations. Traditional fisheries reflect cultural traits and attitudes and may be strongly influenced by religious practices or social customs. Knowledge is transmitted between generations by word of mouth. They are usually small-scale and/or artisanal.

GT: In subsistence fishery fishes or other seafood are harvested for noncommercial, customary and traditional uses. These uses include direct personal or family consumption as food, shelter, fuel, clothing, tools, or transportation, for the making and selling of handicraft articles out of nonedible by-products of fish and wildlife resources taken for personal or family consumption, and for the customary trade, barter, or sharing for personal or family consumption.

In Alaska, subsistence fisheries may not operate in 'nonsubsistence areas' as designated by the state (AS 16.05.258(c)).

A: Subsistence fishery is a type of fishing where the fish caught are shared and consumed directly by the families and kin of the fishers rather than being bought by intermediaries and sold at the next larger market. Pure subsistence fisheries are rare as part of the products are often sold or exchanged for other goods or services. In Alaska, there are no prohibitions on subsistence fisheries. However, there are regulations in place to ensure sustainable management of these resources. For example, there are limits on the number of fish that can be caught per person and seasonal restrictions on when certain species can be harvested. Additionally, there are requirements for proper handling and storage of the catch to prevent waste and spoilage. Overall, subsistence fisheries play an important role in providing food security for many communities in Alaska. By regulating these activities, we can help ensure that they remain sustainable for future generations while also supporting local economies.

5. **Q:** What are the tags implanted in the fish's abdominal cavity?

C: The main types of internal tags are:

passive integrated transponders (PIT tags)

radio transmitters

acoustic transmitters

Internal devices send out a signal that provides a unique code to identify an individual fish. Fish with internal devices are very hard to visually identify because the needle or surgical wounds left after the device is implanted will heal and leave little to no scarring. To address this issue, we often add an external tag to fish along with the internal device

GT: The tags implanted in the fish's abdominal cavity are:

- PIT tags: Passive Integrated Transponder tags are small, passive radio transponder tag which, when in range, are activated by a signal emitted from a tag reader. The tag then emits a unique identification code back to the reader.

- Acoustic tags: small, electronic, sound-emitting devices which collect information on fish's physiology and/or movement patterns. Data are transmitted wirelessly, usually through the use of radio waves, acoustic signals or via satellite communication. Reading is done using the

tag's associated equipment, often through a computer connection.

- Radio tags: emit a radio signal that can be detected by a receiver. Like acoustic tags, radio tags allow researchers to track the movements of tagged fish. The tags are inserted into the fish's abdominal cavity through a needle or surgical wound, which heals and leaves little to no scarring.

A: The tags implanted in the fish's abdominal cavity are passive integrated transponders (PIT tags), radio transmitters, and acoustic transmitters. These tags send out a signal that provides a unique code to identify an individual fish. The tags are inserted into the fish's abdominal cavity through a needle or surgical wound, which heals and leaves little to no scarring.

6. **Q:** Identify and describe the length measurement type used for tuna. What are the automatic methods used for tuna length measurement?

C: Table 2. Methods of fish length measurement.

Method	Object	Measurement error rate	Strength and Weakness
--------	--------	------------------------	-----------------------

Hough transform [3]	Tuna	<5%	+ have relatively high accuracy of measuring tuna length + accuracy rate can be increase by using projective transform
---------------------	------	-----	--

Image thinning [9]	Tuna	<3%	+ have high accuracy on measuring tuna length
--------------------	------	-----	---

Hsiu Method [6]	Fish	<2.19%	+ have high accuracy on measuring smaller fish length - No past research have been done to measure tuna using this method
-----------------	------	--------	---

Best fitting rectangle [2]	Fish	<0.5%	+ have high accuracy on measuring smaller fish length - No past research have been done to measure tuna using this method
----------------------------	------	-------	---

Grade-3 polynomial regression [7]	Fish	<4%	- No past research have been done to measure tuna using this method
-----------------------------------	------	-----	---

3. Discussions

Automation of the tuna length measurement process is important because it will make the length measurement more accurate and faster compare to manual measurement. Thus, it is important to make the measurement process automatic to increase the efficiency of the measurement process and increase the effectiveness of tuna catch regulation. There are other features that need to take into consideration in measuring tuna length using machine vision such as the visual perspective of the object, the angle of the camera to the object, and the

distance between the camera and the object. The angle of the camera is very important as it will greatly affect the visual perspective of the object [3]. There are study had been made by taking in the one of the features for fish length measurement method [6,11] but no research had been done yet by using all of the features for fish length measurement method by considering all of the measurement environment factor. Usage of lightbox have higher accuracy in measuring the length of other smaller fish [2] but the disadvantage of using lightbox is it cannot fit the tuna body as tuna body is too big to be fit into the lightbox. There were a number of researches had been done with smaller fish species, but they were not yet tested on tuna fish [8]. Although the length measurement result has a low error rate, it is still not satisfactory. More research needed to be done in order to increase the accuracy rate of fish length measurement. Based on [2], although the method can be used on both smaller and bigger fish, the result of the length measurement very different from each other.

GT: The length measurement type used for tuna is fork length. Fork length is the distance from the tip of the snout to the fork of the tail. It is the most commonly used length measurement type for tuna. Automatic methods used for tuna length measurement are Hough transform, image thinning, best fitting rectangle, Hsiu method and grade-3 polynomial regression. Each method has its own strengths and weaknesses. For example, the Hough transform has a low error rate of less than 5%, while the Hsiu method has a high accuracy rate of measuring smaller fish lengths. However, these methods have not been widely used for tuna length measurement due to their limitations and challenges.

A: The length measurement type used for tuna is the Hough transform. This method uses a mathematical algorithm to detect lines and curves in images. It is commonly used in computer vision applications, including fish length measurement. To automatically measure the length of tuna, researchers have developed several methods. One popular method is the Hough transform, which is based on the principle of line detection. Another method is image thinning, which involves reducing the thickness of the image to a single pixel width. This makes it easier to identify the edges of the fish and calculate its length. Overall, there are many different methods available for automatically measuring the length of tuna. Each method has its own strengths and weaknesses, so it is important to choose the right one for your specific application.

7. **Q:** Summarize the salmon life-cycle's stages.

C: Salmon Life Cycle

Salmon are native to the world's two biggest oceans and the rivers draining into them. The Atlantic Ocean has only one species, the Atlantic salmon (*Salmo salar*), while in the Pacific Ocean there are several species including Pink (*Oncorhynchus gorbuscha*), Chum (*O. keta*), Sockeye (*O. nerka*), Coho (*O. kisutch*), Chinook (*O. tshawytscha*) and Amago (*O. rhodurus*).

Life Cycle of Atlantic Salmon*

Irish salmon are Atlantic salmon and spend their juvenile phase in rivers before migrating to sea to grow and mature. To complete their life cycle they must return to their river of origin to spawn. The salmon who adopt this life cycle are called anadromous.

Life begins*

The salmon starts life as a small pea sized egg hidden away under loose gravel in cool clean rivers entering the North Atlantic Ocean. Against the odds the parents of this little egg have succeeded in returning to freshwater to spawn completing their life cycle before giving rise to a new generation. To do this both male and female adults ceased to feed on entering freshwater in response to gonadal development, directing all their energy instead to reproduction. The migration of adults in winter to suitable habitat can commence up to a year before spawning takes place. Spawning typically occurs in headwaters, though it may happen anywhere in a river if a suitable substrate of well oxygenated loose gravel is available. At spawning time (November to January), the female digs a depression in the gravel with her tail to deposit her eggs. One or more males discharge milt over the falling eggs to fertilize. Quickly the female covers the eggs with gravel to a depth of several centimetres which forms a nest or "redd" on the river bed. Buried deep inside the gravel the ova are safe from the impact of debris carried along in heavy floods and from attack by predators such as eels (*Anguilla anguilla*), trout (*Salmo trutta*) or cormorants (*Phalacrocorax carbo*).

Ova*

The rate of egg or "ova" development is dependent on water temperature. Eyes inside the pea sized orange ova are visible and increasing movement can be detected as the yolk sac containing food is consumed. The number of ova deposited in the redd is determined by the

size of the female with larger females over 10kg depositing 15,000 each. This high fecundity (ova per female) is critical as survival in the wild is extremely low, especially in freshwater. For example, in the Burrishoole River on the west coast of Ireland survival rates for juveniles from 1970 to 2015 was as low as 0.3% rising to a high of just 1.3% in 2007.

Alevins*

The just-hatched fish are called "alevins" and still have the yolk sac attached to their bodies in Spring. When their yolk sac is absorbed the alevins become increasingly active and begin their journey up through the gravel of the riverbed. When strong enough the small fish must rise to the surface of the water and gulp air. By doing this they fill their swim bladder to gain neutral buoyancy making it easier to swim and hold their position in fast flowing streams. This critical period is therefore referred to as "swim- up" and exposes the young to dangerous predators for the first time. Once they begin to swim freely they are called fry.

Fry*

The fry have eight fins, which are used to maintain their position in fast flowing streams and manoeuvre about in the water during the Summer months. Fry feed on microscopic invertebrates and their abundance is regulated by temperature, predation, pollution and competition for food with other fry and other species of fish. The presence of salmon in a river is synonymous with a healthy aquatic environment, and as they are extremely sensitive to changes in water quality, habitat and climate, salmon are a good indicator of freshwater and marine ecosystem status.

Parr*

Over the Autumn the fry develop into parr with vertical stripes and spots for camouflage. They feed on aquatic insects and continue to grow for one to three years while maintaining their territory in the stream. Once the parr have grown to between 10 and 25cm in body length, they undergo a physiological pre- adaptation to life in seawater by smolting. This is evident by changes in their appearance as they become silvery and swim with the current instead of against it. There are also internal changes in the salt-regulating mechanisms of the fish. This adaptation prepares the smolt for its journey to the ocean.

Smolt*

In Spring, large numbers of smolts 1-3 years old, leave Irish rivers to migrate along the North

Atlantic Drift, and into the rich feeding grounds of the Norwegian Sea and the greater expanse of the North Atlantic Ocean. Here they feed primarily on fish such as capelin (*Mallotus villosus*) , herring (*Alosa* spp.), and sand eel (*Ammodytes* spp.). As they grow quickly fewer predators are able to feed on them. Their rate of growth is therefore critical to survival.

Adult Salmon*

Salmon that reach maturity after one year at sea are called Grilse; these return to their river in summer weighing from 0.8 to 4kg. If it takes two or more years at sea to mature the salmon will return considerably earlier in the year and larger at 3 to 15kg, and because of their size they are greatly sought after by fishermen. Salmon exhibit remarkable "homing instinct" with a very high proportion able to locate their river of origin using the earth's magnetic field, the chemical smell of their river and pheromones (chemical substances released by other salmon in the river). Perfect homing precision is expected even after migrations over 3,000km to feeding grounds north of the Arctic Circle in the Norwegian Sea and at West Greenland. There is great excitement when adult salmon return to rivers as many are seen leaping acrobatically into the air and jumping over waterfalls while moving upstream. Salmon that survive fishermen, poachers and pollution may still have to scale large dams built across rivers before eventually finding refuge in lakes and deep pools. Arriving upstream on their spawning grounds among big boulders in icy headwaters the life cycle begins again, so ensuring survival of the species for another generation.

Kelts*

Having spawned, the salmon are referred to as "kelts". Weak from not eating since arriving in freshwater and losing energy in a bid to reproduce successfully, kelts are susceptible to disease and predators. Mortality after spawning can be significant, especially for males but some do survive and commence their epic journey again. Scientists studying salmon initially used the rings laid down on scales, much like tree rings, to determine the age and growth of salmon in freshwater and at sea. By doing this they established that some kelts succeeded in spawning three times! Now a new record exists of an Irish salmon that reached maturity after less than one year at sea - a zero sea winter salmon

GT: The life stages of salmon are summarized as follows:

Egg (Ova): the salmon begins life as a pea-sized egg, hidden under loose gravel in cool, clean

ivers. Eggs have a high mortality rate, with only a small percentage surviving to hatch.

Alevin: upon hatching in spring, the fish are called alevins and have a yolk sac attached, providing nourishment. Once the yolk sac is absorbed, alevins become active, move through the gravel, and must gulp air to fill their swim bladders for neutral buoyancy.

Fry: they have eight fins and feed on microscopic invertebrates during summer; they are typically found in shallow waters near the shoreline.

Parr: in autumn, fry develop into parr; they have vertical stripes and spots for camouflage and they feed on aquatic insects, growing for 1-3 years while establishing territories in the stream.

Juvenile: young fish, mostly similar in form to adult but not yet sexually mature.

Smolt: once they reach 10-25 cm, parr undergo smolting; they become silvery and start swimming with the current, preparing for ocean migration.

Jack: precocial male salmon that have spent one winter less in the ocean than the youngest females of a given species.

Grilse (Adult Salmon): smolts migrated to the ocean that exhibit strong homing instincts to return to their river of origin to spawn.

Kelt: after spawning, salmon are referred to as kelts; weakened from not eating and the energy expended in reproduction, many kelts die.

A: The salmon life cycle consists of various stages, including egg, alevin, fry, parr, smolt, and adult salmon. The salmon begins its life cycle as an egg, which is laid in the gravel of a river. After hatching, the alevin stage begins, where the young salmon absorbs its yolk sac and begins swimming upstream. During the fry stage, the salmon feeds on small invertebrates and grows until it reaches the parr stage. In the parr stage, the salmon develops vertical stripes and spots for camouflage and continues to grow until it reaches the smolt stage. The smolt stage is characterized by the salmon's ability to adapt to living in saltwater, and it migrates to the ocean to feed and grow. After spending one to three years in the ocean, the salmon returns to its home river to spawn, completing the life cycle.

8. **Q:** List all pacific salmon species. Which is the most widespread species?

C: Pacific salmon leaping at Willamette Falls, Oregon Commercial production of salmon in

million tonnes 1950–2010[44] Life cycle of Pacific salmon

Some other fishes called salmon

Common name	Scientific name	Order	Maximum length	Common length	Maximum weight	Maximum age	Trophic level	Fish Base	FAO	ITIS	IUCN status
Australian salmon	<i>Arripis trutta</i> (Forster, 1801)	Perciformes	89 cm (2 ft 11 in)	47 cm (1 ft 7 in)	9.4 kilograms (21 lb)	26 years	4.1	[34]	[35]	Not assessed	
Danube salmon	<i>Hucho hucho</i> (Linnaeus, 1758)	Salmoniformes	150 cm (4 ft 11 in)	70 cm (2 ft 4 in)	52 kilograms (115 lb)	15 years	4.2	[36]	[37]	Endangered	[38]
Hawaiian salmon	<i>Elagatis bipinnulata</i> (Quoy & Gaimard, 1825)	Carangiformes	180 cm (5 ft 11 in)	90 cm (2 ft 11 in)	46.2 kilograms (102 lb)	6 years	3.6	[39]	[40]	[41]	Not assessed
Indian salmon	<i>Eleutheronema tetradactylum</i> (Shaw, 1804)	Perciformes	200 cm (6 ft 7 in)	50 cm (1 ft 8 in)	145 kilograms (320 lb)	years 4.4	[42]	[43]	Not assessed		

Atlantic salmon (*Salmo salar*) reproduce in northern rivers on both coasts of the Atlantic Ocean. Landlocked Atlantic salmon (*Salmo salar* m. *sebago*) is a potamodromous (migratory only between fresh waters) subspecies/morph that live in a number of lakes in eastern North America and in Northern Europe, for instance in lakes Sebago, Onega, Ladoga, Saimaa, Vänern and Winnepesaukee. They are not a different species from the sea-run Atlantic salmon but have independently evolved a freshwater-only life cycle, which they maintain even when they could access the ocean. Chinook salmon (*Oncorhynchus tshawytscha*) are also known in the United States as king salmon or "blackmouth salmon", and as "spring salmon" in British Columbia, Canada. Chinook salmon is the largest of all Pacific salmon, frequently exceeding 6 ft (1.8 m) and 14 kg (30 lb).[45] The name tyee is also used in British Columbia to refer to Chinook salmon over 30 pounds and in the Columbia River watershed, especially large Chinooks were once referred to as June hogs. Chinook salmon are known to range as far north as the Mackenzie River and Kugluktuk in the central Canadian arctic,[46] and as far south as the Central Californian Coast.[47] Chum salmon (*Oncorhynchus keta*) is known as dog salmon or calico salmon in some parts of the US, and as keta in the Russian Far East. This species has the widest geographic range of the Pacific species:[48] in the eastern Pacific from north of the Mackenzie River in Canada to south of the Sacramento River in California and in the western Pacific from Lena River in Siberia to the island of Kyūshū in the Sea of Japan. Coho salmon (*Oncorhynchus kisutch*) are also known in the US as silver salmon. This species is found throughout the coastal waters of Alaska and British Columbia and as far south as Central California (Monterey Bay).[49] It is

also now known to occur, albeit infrequently, in the Mackenzie River.[46] Masu salmon (*Oncorhynchus masou*), also known as "cherry trout" (, sakura masu) in Japan, are found only in the western Pacific Ocean in Japan, Korea, and Russian Far East. A landlocked subspecies known as the Taiwanese salmon or Formosan salmon (*Oncorhynchus masou formosanus*) is found in central Taiwan's Chi Chia Wan Stream.[50] Pink salmon (*Oncorhynchus gorbuscha*), known as humpback salmon or "humpies" in southeast and southwest Alaska, are found in the western Pacific from Lena River in Siberia to Korea, found throughout northern Pacific, and in the eastern Pacific from the Mackenzie River in Canada[46] to northern California, usually in shorter coastal streams. It is the smallest of the Pacific species, with an average weight of 1.6 to 1.8 kg (3.5 to 4.0 lb).[51] Sockeye salmon (*Oncorhynchus nerka*) is also known as red salmon in the US (especially Alaska).[52] This lake-rearing species is found in the eastern Pacific from Bathurst Inlet in the Canadian Arctic to Klamath River in California, and in the western Pacific from the Anadyr River in Siberia to northern Hokkaidō island in Japan

GT: The Pacific salmon species are:

- Chinook salmon (*Oncorhynchus tshawytscha*): Also known as king salmon or blackmouth salmon, and spring salmon in British Columbia.
- Chum salmon (*Oncorhynchus keta*): Known as dog salmon or calico salmon in some parts of the US, and keta in the Russian Far East.
- Coho salmon (*Oncorhynchus kisutch*): Also known as silver salmon.
- Masu salmon (*Oncorhynchus masou*): Also known as cherry trout in Japan.
- Pink salmon (*Oncorhynchus gorbuscha*): Known as humpback salmon or humpies in southeast and southwest Alaska.
- Sockeye salmon (*Oncorhynchus nerka*): Also known as red salmon in the US, especially Alaska.

Chinook salmon is the most widespread species.

A: The Pacific salmon species include: 1. Chinook salmon (*Oncorhynchus tshawytscha*) 2. Coho salmon (*Oncorhynchus kisutch*) 3. Sockeye salmon (*Oncorhynchus nerka*) 4. Pink salmon (*Oncorhynchus gorbuscha*) 5. Chum salmon (*Oncorhynchus keta*) 6. Masu salmon (*Oncorhynchus masou*) 7. Landlocked Atlantic salmon (*Salmo salar* m. sebago) Among these species, Chinook salmon is the most widespread, with a range extending from the Mackenzie

River in Canada to the Central Californian Coast.

9. **Q:** What is aquaculture? How important is it for fish human consumption?

C: Aquaculture has expanded fish availability to regions and countries with otherwise limited or no access to the cultured species . At the global level, since 2016, aquaculture has been the main source of fish available for human consumption. In 2023, this share was 56% and is expected to increase in the long term. In 2023, aquaculture accounted for 90 million tonnes (LW) destined for direct human food consumption, while wild capture accounted for 7.2 million tonnes (LW). However, fish has been estimated to account for only 7% of global protein consumption (and about 17% of total fish and animal protein supply). Aquaculture production in tonnes of all true salmon species 1950–2010, as reported by the FAO[44] Salmon farming sea cage in Torskefjorden, Senja Island, Troms, Norway Artificially incubated chum salmon fries Salmon aquaculture is a major contributor to the world production of farmed finfish, representing about US\$10 billion annually

GT: Aquaculture is the culturing of fish, shellfish, aquatic plants, and/or other organisms in captivity or under controlled conditions in the near shore environment. It is an important source of fish for human consumption, providing about half of the world's fish supply. In fact, in 2023, aquaculture accounted for 90 million tonnes (LW) of fish destined for direct human food consumption. Furthermore, aquaculture is a major contributor to the world's economy, with an annual value of around \$10 billion.

A: Aquaculture is the farming of fish, shellfish, and other aquatic organisms. It is an important source of food for humans, providing a significant portion of the world's fish supply. In 2023, aquaculture accounted for 90 million tonnes (LW) of fish destined for direct human food consumption, while wild capture accounted for 7.2 million tonnes (LW). Fish has been estimated to account for only 7% of global protein consumption (and about 17% of total fish and animal protein supply). Aquaculture is also important for the economy, providing jobs and income for many people around the world.

10. **Q:** What are the different regulations for fish farming in Canada based on geographical area?

C: Barriers to Entry - Licenses 91 | Page Regulation of fish farming in Canada License and location Fish farming companies in Canada are subject to different regulations depending on the geographical area they operate in. The principal Federal laws are the Fisheries Act, the Canadian Navigable Waters Act, The Health of Animals Act, and the Species at Risk Act. The Aquaculture Activities Regulations (AAR) are national regulations that apply throughout Canada. Each province has specific Acts and Regulations that also apply. The three primary fish farming areas in Canada are British Columbia, Newfoundland and Labrador, and New Brunswick. To operate a marine fish farm site, provincial and/or federal authorisations are required. In Newfoundland and Labrador and New Brunswick, the Provincial government is the primary regulator and leasing authority. The Province regulates the activity and operations of aquaculture and issues the Aquaculture License, Crown Land lease and Water Use License where fish farms are located. In Newfoundland and Labrador, the Crown Land Lease for the site is issued for 50 years, the Aquaculture License is issued for 6 years, and the Water Use License is issued for 5 years. In New Brunswick, individual sites are typically granted lease for 20 years. All Commercial Aquaculture Licenses are renewable but may be lost or suspended for non-compliance issues and non-payment of fees. Farms in New Brunswick are also issued an Approval to Operate with a 5 year term limit by the provincial department of Environment Department. In British Columbia, Federal and Provincial authorisations are required to operate a marine fish farm site. The Federal Government regulates the activity and operations of aquaculture while the Provincial Government administers the Crown lands where fish farms are located. The Province grants a license to occupy an area of the ocean associated with the individual fish farming site. The tenure encompasses the rearing pens, ancillary infrastructure and all moorings. Individual site tenures have a specific timeline ranging from five to twenty years. The term of tenure is based upon the provincial policy at the time of offer. In 2024, the annual fee for a typical 35 hectares tenure is \$22,000 CAD per year. This fee is calculated based on the tenure size and a provincially indexed land value. Each tenure license contains a renewal provision once expired. After the tenure term has expired, it becomes a month to month occupancy until it is either renewed or returned to the Crown. It is uncommon for a tenure to not be renewed, however breaches to a tenure agreement can result in non-renewal. The production limitations in Canada are regulated

as either a “Maximum Allowable Biomass” or a fixed number of smolt per cycle. “MAB” is specific to each Aquaculture licensed facility in British Columbia. Smaller farms are typically licensed for 2,200mt. with larger capacity facilities licensed to produce 5,000 mt. per cycle. In Newfoundland and Labrador and New Brunswick, a maximum number of smolt per cycle is given to a farm. Farms are typically licensed for 600,000 to 1,000,000 smolt per cycle in Newfoundland and Labrador, and 250,000 to 600,000 smolt per cycle in New Brunswick. In British Columbia, the Federal Government grants an Aquaculture License with conditions that a farm must meet. The Aquaculture license conditions are linked to The Fisheries Act. Aquaculture license conditions specify the species being farmed, the Maximum Allowable Biomass (MAB) on the site, the type of rearing equipment and the allowable environmental impact. Production or “MAB” is specific to each site. The annual license fee is calculated at \$2.95 CAD per ton of MAB for operational sites

GT: Fish farming companies in Canada are subject to different regulations depending on the geographical area they operate in. The three primary fish farming areas in Canada are British Columbia, Newfoundland and Labrador, and New Brunswick. In Newfoundland and Labrador and New Brunswick, the Provincial government is the primary regulator and leasing authority. The Province regulates the activity and operations of aquaculture and issues the Aquaculture License, Crown Land lease, and Water Use License where fish farms are located. In Newfoundland and Labrador, the Crown Land Lease for the site is issued for 50 years, the Aquaculture License is issued for 6 years, and the Water Use License is issued for 5 years. In New Brunswick, individual sites are typically granted a lease for 20 years. In British Columbia, Federal and Provincial authorizations are required to operate a marine fish farm site. The Federal Government regulates the activity and operations of aquaculture while the Provincial Government administers the Crown lands where fish farms are located. The Province grants a license to occupy an area of the ocean associated with the individual fish farming site. The tenure encompasses the rearing pens, ancillary infrastructure, and all moorings. Individual site tenures have a specific timeline ranging from five to twenty years.

A: There are several regulations for fish farming in Canada based on geographical area. In British Columbia, the federal government regulates the activity and operations of aquaculture, while the provincial government administers the crown lands where fish farms are located.

The provincial government grants a license to occupy an area of the ocean associated with the individual fish farming site. The tenure encompasses the rearing pens, ancillary infrastructure, and all moorings. The term of tenure is based upon the provincial policy at the time of offer. In 2024, the annual fee for a typical 35 hectares tenure is \$2,200 CAD per year. This fee is calculated based on the tenure size and a provincially indexed land value. Each tenure license contains a renewal provision once expired. After the tenure term has expired, it becomes a month to month occupancy until it is either renewed or returned to the Crown. It is uncommon for a tenure to not be renewed, however breaches to a tenure agreement can result in non-renewal. The production limitations in Canada are regulated as either a "Maximum Allowable Biomass" or a fixed number of smolt per cycle. "MAB" is specific to each Aquaculture license in British Columbia. Smaller farms are typically licensed for 2,200mt. with larger capacity facilities licensed to produce 5,000 mt. per cycle. In Newfoundland and Labrador and New Brunswick, a maximum number of smolt per cycle is given to a farm. Farms are typically licensed for 600,000 to 1,000,000 smolt per cycle in Newfoundland and Labrador, and 250,000 to 600,000 smolt per cycle in New Brunswick. In Newfoundland and Labrador and New Brunswick, the provincial government is the primary regulator and leasing authority. The Province regulates the activity and operations of aquaculture and issues the Aquaculture License, Crown Land lease, and Water Use License where fish farms are located. In Newfoundland and Labrador, the Crown Land Lease for the site is issued for 50 years, the Aquaculture License is issued for six years, and the Water Use License is issued for five years. All Commercial Aquaculture Licenses are renewable but may be lost or suspended for non-compliance issues and non-payment of fees. Farms in New Brunswick are also issued an Approval to Operate with a five-year term limit by the provincial Department of Environment.

11. **Q:** What are the key indicators for projecting future fish harvest volumes? How does smolt release data affect long-term volume estimates? How does seawater temperature affect production cycle length and harvest volumes? How do disease outbreaks impact fish harvest volumes?

C: Indicators Determining Harvest volumes 104 | Page Projecting future harvest volumes
The three most important indicators for future harvest volumes are standing biomass, feed

consumption and smolt release. These are good indicators for medium - and long -term harvest, while the best short -term indicator is standing biomass categorized by size. As harvested size is normally above 4 kg, the available biomass of this size class is therefore the best estimate of short -term supply. If no actual numbers on smolt releases are available, vaccine sales could be a good indicator of number of smolt releases and when the smolt is put to sea. This is a good indicator of long -term harvest volumes as it takes up to 2 years from smolt release to harvest. Variation in seawater temperature can materially impact the length of the production cycle. A warmer winter can for example increase harvest volumes for the relevant year, partly at the expense of the subsequent year. Disease outbreaks can also impact harvest volume due to mortality and growth slowdown3% of global seafood supply 14 Considerable opportunities within aquaculture 15 Supply of farmed and wild salmonids 16 Salmonids harvest 2023 17 3 SALMON DEMAND 18 Global macro trends 19 Favourable attributes of salmon 20 Relative price development of protein products 24 4 SALMON SUPPLY 25 Total harvest of Atlantic salmon 2003 -2023 26 Diminishing growth expectations 27 Few coastlines suitable for salmon farming 28 5 SUSTAINABLE PRODUCTION 29 UN's Sustainable Development Goals 30 Environmental impact of aquaculture 31 Material sustainability efforts 34 Sustainability of fish feed 35 Global sustainability initiatives 37 Transparency 38 6 SALMON MARKETS 39 Global trade flow of farmed Atlantic salmon 40 Farmed Atlantic salmon by market 41 Top 10 markets by size (2023E) 42 Development of value (CAGR 8%) vs. volume last 10 years 43 Price neutral demand growth – approx

GT: The three key indicators for projecting future fish harvest volumes are standing biomass, feed consumption, and smolt release. Standing biomass categorized by size is the best short-term indicator, while standing biomass, feed consumption, and smolt release are good indicators for medium- and long-term harvest. Smolt release data affects long-term volume estimates as it takes up to 2 years from smolt release to harvest. Variation in seawater temperature can impact the length of the production cycle and harvest volumes. A warmer winter can increase harvest volumes for the relevant year, partly at the expense of the subsequent year. Disease outbreaks can impact harvest volume due to mortality and growth slowdown.

A: The key indicators for projecting future fish harvest volumes include standing biomass, feed consumption, and smolt release. Smolt release data affects long-term volume estimates

by providing information about the number of fish released into the ocean, which can help predict future harvests. Seawater temperature can also impact production cycle length and harvest volumes, as warmer water can lead to faster growth rates and shorter production cycles. Disease outbreaks can impact fish harvest volumes by causing mortality and slowing down growth rates.

12. **Q:** How do husbandry and health practices contribute to maximise salmon survival and fish stock maintenance?

C: Risk Factors 99 | Page Salmon health and welfare Maximising survival and maintaining healthy fish stocks are primarily achieved through good husbandry and health management practices and policies, which reduce exposure to pathogens and the risk of health challenges. The success of good health management practices has been demonstrated on many occasions and has contributed to an overall improvement in the survival of farmed salmonids. Fish health management plans, veterinary health plans, biosecurity plans, risk mitigation plans, contingency plans, disinfection procedures, surveillance schemes , as well as coordinated and synchronised zone/area management approaches, all support healthy stocks with an emphasis on disease prevention. Prevention of many diseases is achieved through vaccination at an early stage and while the salmon are in freshwater. Vaccines are widely used commercially to reduce the risk of health challenges. With the introduction of vaccines a considerable number of bacterial and viral health issues have been effectively controlled, with the additional benefit that the quantity of licensed medicines prescribed in the industry has been reduced. In some instances medicinal treatment is still required to avoid mortality and for the well-being and welfare of the fish Salmon Farming Industry Handbook 2024

GT: Maximising survival and maintaining healthy fish stocks are primarily achieved through good husbandry and health management practices and policies, which reduce exposure to pathogens and the risk of health challenges. The success of good health management practices has been demonstrated on many occasions and has contributed to an overall improvement in the survival of farmed salmonids. Fish health management plans, veterinary health plans, biosecurity plans, risk mitigation plans, contingency plans, disinfection procedures, surveillance schemes, as well as coordinated and synchronised zone/area management approaches,

all support healthy stocks with an emphasis on disease prevention. Prevention of many diseases is achieved through vaccination at an early stage and while the salmon are in freshwater. Vaccines are widely used commercially to reduce the risk of health challenges. With the introduction of vaccines a considerable number of bacterial and viral health issues have been effectively controlled, with the additional benefit that the quantity of licensed medicines prescribed in the industry has been reduced. In some instances medicinal treatment is still required to avoid mortality and for the well-being and welfare of the fish. Even the best managed farms may have to use licensed medicines from time to time, if other measures are not sufficient.

A: Husbandry and health practices play a crucial role in maximizing salmon survival and maintaining healthy fish stocks. Good husbandry practices include proper feeding, water quality management, and disease prevention measures. Health practices involve regular monitoring of fish health, timely diagnosis and treatment of diseases, and implementation of preventive measures such as vaccinations and biosecurity protocols. These practices help to minimize stress on the fish, reduce the risk of disease outbreaks, and ensure optimal growth and survival rates. By implementing effective husbandry and health practices, salmon farmers can maintain healthy fish stocks and maximize their production efficiency.

13. **Q:** Why is *salmo salar* considered a healthy product?

C: There is a rising demand for more sustainable food and a willingness to pay for it. The sustainable properties of salmon therefore make the product attractive to consumers.

GT: Atlantic salmon is a healthy product because is rich in long-chain omega-3, EPA and DHA, which reduce the risk of cardiovascular disease. Data also indicates that EPA and DHA reduce the risk of a large number of other health issues. Furthermore, it's nutritious, rich in micronutrients, minerals, marine omega-3 fatty acids, high-quality protein and several vitamins.

A: There is a rising demand for more sustainable food and a willingness to pay for it. The sustainable properties of salmon therefore make the product attractive to consumers.

14. **Q:** What are the main fish stock types?

C: * Search

Types of fisheries

Capture fisheries are extremely diversified, comprising a large number of types of fisheries that are categorized by different levels of classification. On a broad level, capture fisheries can be classified as ****industrial, small-scale/artisanal and recreational****. A more specific level includes reference to the fishing area, gear and the main target species, such as the North Sea herring purse seine fishery, Gulf of Mexico shrimp trawl fishery, southern ocean Patagonian toothfish longline fishery. While capture fisheries encompass thousands of fisheries on a global scale, they are often categorized by the capture species, the fishing gear used and the level at which a fishery is managed nationally and/or regionally.

The following brief descriptions provide an overview of capture fishery types.

Industrial fisheries

Capital-intensive fisheries using relatively large vessels with a high degree of mechanization and that normally have advanced fish finding and navigational equipment. Such fisheries have a high production capacity and the catch per unit effort is normally relatively high. In some areas of the world, the term "industrial fisheries" is synonymous with fisheries for species that are used for reduction to fishmeal and fish oil (e.g. the trawl fishery for sandeel in the North Sea or the Peruvian ouse-seine fishery for anchoveta).

Small-scale fisheries

Labour-intensive fisheries using relatively small crafts (if any) and little capital and equipment per person-on-board. Most often family-owned

GT: The main fish stock types are:

- Aquaculture: the culturing of fish, shellfish, aquatic plants, and/or other organisms in captivity or under controlled conditions in the near shore environment.
- Hatchery: the artificial breeding, hatching, and rearing through the early life stages of animals – finfish and shellfish in particular.
- Wild stock: a stock that is sustained by natural spawning and rearing in the natural habitat, regardless of parentage or origin.

A: The main fish stock types include industrial fisheries, small-scale fisheries, and recreational

fisheries. Industrial fisheries use large vessels with advanced technology, while small-scale fisheries rely on smaller boats and manual labor. Recreational fisheries involve anglers who fish for sport or leisure.

C Salmon RAGOnto queries

1. **Question:** Give a definition of seine net. What is the purpose of pair seines in fishing?

These concepts might be useful to answer the question:

seine, beach seine, purse seine, netting, fishing gear type

Context: Seines

A seine net is a very long net, with or without a bag in the centre, which is set either from the shore or from a boat for surrounding a certain area and is operated with two (long) ropes fixed to its ends (for hauling and herding the fish) Pair seines (SPR)

A seine net designed for pair fishing, where two boats operate the net in pair

Ground Truth: the same as the salmon RAG one

Answer: A seine net is a very long net, with or without a bag in the centre, which is set either from the shore or from a boat for surrounding a certain area and is operated with two (long) ropes fixed to its ends (for hauling and herding the fish). Pair seines (SPR) are seine nets designed for pair fishing, where two boats operate the net in pair.

2. **Q:** What fishing net prevents fish from escaping using netting panels?

These concepts might be useful to answer the question:

netting, gillnet, set gillnet, drift gillnet, seine, beach seine, purse seine, fishing gear type

C: Surrounding nets and lift nets

Surrounding nets are large netting walls set for surrounding aggregated fish both from the sides and from underneath, thus preventing them from escaping by diving downwards

GT: the same as the salmon RAG one

A: Surrounding nets and lift nets

3. **Q:** What are trollers in commercial salmon fisheries?

These concepts might be useful to answer the question:

troll, fishing gear type

C: The troll fishery generally occurs July 1, and remains open until a set number of fish are caught. Licence conditions and commercial fishing plans lay out allowable gear characteristics such as hook styles, mesh size, net dimensions and the methods by which gear may be used (e.g. set times for nets, mandatory brailing and sorting of fish). On the North Coast, the commercial net fishery is open in defined terminal areas of various systems, notably the Skeena/Nass systems and the Bella Coola/Atnarko. Openings could occur anywhere inside the surfline depending on local stock strengths. Please see our glossary of salmon terms for more information.

troll: One or more fishing lines, which are baited with lures or bait fish and drawn through the water (often behind a moving vessel) to catch pelagic fishes. fishing gear type: Equipment which is used to harvest aquatic resources, and in particular fish(es). fishing gear type: [ADDRESS THIS COMMENT, THEN DELETE]

Could use some input on subclass hierarchy. E.g. 'Hand collection' includes methods which are largely done with tools that can be operated by hand. 'Netting' includes things like gillnets and seines, which, while largely set from fishing vessels (particularly in commercial fisheries), can also be set by hand at smaller scales.

GT: the same as the salmon RAG one

A: Trollers are a type of fishing gear used in commercial salmon fisheries. They consist of one or more fishing lines, which are baited with lures or bait fish and drawn through the water (often behind a moving vessel) to catch pelagic fishes. Trolling is often used in conjunction with other fishing techniques, such as bottom trawling or longlining, to increase the chances of catching fish.

4. **Q:** What is subsistence fishery? Is there any prohibition in Alaska?

These concepts might be useful to answer the question:

subsistence fishery, fishery type

C: subsistence fishery: In Alaska, subsistence fisheries may not operate in "nonsubsistence

areas" as designated by the state (AS 16.05.258(c)). subsistence fishery: A fishery in which fishes or other seafood are harvested for noncommercial, customary and traditional uses. These uses include direct personal or family consumption as food, shelter, fuel, clothing, tools, or transportation, for the making and selling of handicraft articles out of nonedible by-products of fish and wildlife resources taken for personal or family consumption, and for the customary trade, barter, or sharing for personal or family consumption. fishery type: The Food and Agriculture Organization of the United Nations (FAO) defines a fishery as ".an activity leading to harvesting of fish. It may involve capture of wild fish or raising of fish through aquaculture. **GT**: the same as the salmon RAG one

A: Subsistence fishery is a type of fishing that is done for personal or family use rather than for commercial purposes. In Alaska, subsistence fisheries may not operate in "nonsubsistence areas" as designated by the state (AS 16.05.258(c)). Subsistence fishery is a fishery in which fishes or other seafood are harvested for noncommercial, customary and traditional uses. These uses include direct personal or family consumption as food, shelter, fuel, clothing, tools, or transportation, for the making and selling of handicraft articles out of nonedible by-products of fish and wildlife resources taken for personal or family consumption, and for the customary trade, barter, or sharing for personal or family consumption. Fishery type is defined by the Food and Agriculture Organization of the United Nations (FAO) as ".an activity leading to harvesting of fish. It may involve capture of wild fish or raising of fish through aquaculture.

5. **Q**: What are the tags implanted in the fish's abdominal cavity?

These concepts might be useful to answer the question:

fish tag code or color, identifiers and codes, fish tag or identification type, subcutaneous tag, coded wire tag, subcutaneous color mark, visible implant elastomer, tattoo, genetic identification, dye tag, bismarck brown dye tag, calcein dye tag, fish body modification, fin or maxilla clip, pectoral fin clip, left pectoral fin clip, right pectoral fin clip, adipose fin clip, lower caudal fin clip, upper caudal fin clip, maxilla clip, pelvic fin clip, freeze brand, otolith marking, electronic tag, acoustic tag, internal acoustic tag, external acoustic tag, rfid, pit tag, external visual tag, carlin tag, dart tag, natural pattern or unique marking

C: Fish with internal devices are very hard to visually identify because the needle or surgical wounds left after the device is implanted will heal and leave little to no scarring. To address this issue, we often add an external tag to fish along with the internal device.

Report a fish tag

To help us with our research, you can report a tagged fish you have caught. When you report a fish tag, we will ask you to provide the following information:

when and where you caught the tagged fish

the fish's unique tag identification number

whether you harvested (kept) or released the fish

This important information allows our biologists and researchers to better understand our fish populations and develop science-based policy, fishing regulations and conservation initiatives.* * *

Overview

A fish tag is a marking or device that helps biologists and researchers identify an individual fish or groups of fish. Fish tags can be located either:

on the outside of the fish (known as external tags)

inside the body of the fish (known as internal tags)

Why we tag fish

We tag fish to collect information about fish:

population size

mortality rates

growth

movement

habitat use

To learn about the fish population using external tags, we require other information from the person who caught the fish, such as:

the fish species

when and where you caught the tagged fish

the fish's unique tag identification number

whether you harvested (kept) or released the fish

If we know how many tagged fish are captured and harvested by anglers, biologists can estimate **population size and** mortality rate of a fish population (how many fish died).

Many internal fish tag devices actively transmit a signal. Biologists or researchers can pick up the signal by tracking fish in the field or by using stationary receivers located in the environment near the fish (Figure 1). The data from the signals can provide more detailed information on fish movement and habitat use, which is difficult to get from external tags.

Image

Figure 1: Examples of a) a biologist actively tracking fish movement using a radio antenna and b) a stationary acoustic receiver on the lake bottom.

We have tagged many fish species in watersheds across Ontario. Learn about our ongoing fish tagging studies on Ontario GeoHub:

Lake Ontario Atlantic salmon

External fish tags

External tags are often made of metal or plastic. They can be clamped onto the body or inserted into the fish's muscle tissue, like a piercing.

Some of the most common external tags we use are: * jaw tags

spaghetti tags

loop tags

disc tags

All external fish tags have a unique code that allows us to identify the fish. Other information printed on tags include the name of the agency or institution that tagged the fish and their contact information so they can be notified if a fish is caught.

Image

Figure 2: Examples of fish tags we use a) jaw tag, b) spaghetti tag, c) loop tag, and d) disc tag.

Internal tags

Image

Figure 3: Examples of internal tags used by the MRNF: a) Embedded Passive Transponder, b) radio transmitter, c) acoustic transmitter."

Internal devices are used to:

identify a fish

provide information on fish movement and habitat use

Internal devices can be implanted with a needle or surgically implanted in the fish's abdominal cavity. The main types of internal tags are:

passive integrated transponders (PIT tags)

radio transmitters

acoustic transmitters

Internal devices send out a signal that provides a unique code to identify an individual fish

GT: the same as the salmon RAG one

A: The tags implanted in the fish's abdominal cavity are usually small, round, and made of either metal or plastic. These tags are typically embedded under the skin and are not visible to the naked eye. Some examples of tags used for fish identification include:

- PIT tags: Passive Integrated Transponder tags are small, electronic chips that are implanted under the skin of the fish. These tags emit a unique code that can be read by a scanner, allowing researchers to identify individual fish.

- Acoustic tags: These tags are similar to PIT tags but emit a sound signal instead of a code. Researchers can detect these signals using specialized equipment, allowing them to track the movements of tagged fish.

- Radio tags: These tags emit a radio signal that can be detected by a receiver. Like acoustic tags, radio tags allow researchers to track the movements of tagged fish.

- Visual tags: These tags are attached to the outside of the fish and are visible to the naked eye. They may be made of metal, plastic, or other materials and can be used to identify individual fish.

It's important to note that the specific type of tag used for fish identification can vary depending on the species being studied and the purpose of the study. For example, some tags may be designed to last longer than others, while others may be more suitable for tracking fish in certain environments. In conclusion, the tags implanted in the fish's abdominal cavity are typically small, electronic devices that are used to identify individual fish. These tags can be made of metal or plastic and may be embedded under the skin or attached to the outside

of the fish. The specific type of tag used can vary depending on the species being studied and the purpose of the study.

6. **Q:** Identify and describe the length measurement type used for tuna. What are the automatic methods used for tuna length measurement?

These concepts might be useful to answer the question:

fish length measurement type, fork length, fork length measurement method, fish length determination method, fish measurement method, measurement method, post-orbit to fork of tail length, post-orbit to fork of tail measurement method, post orbit to hypural plate length, post-orbit to hypural plate measurement method, mid-orbit to hypural plate length, mid-orbit to hypural plate measurement method, mid-orbit to posterior insertion of anal fin length, mid-orbit to posterior insertion of anal fin measurement method, mid-orbit to fork of tail length, mid-orbit to fork of tail measurement method, total length, total length measurement method, standard length, standard length measurement method, fish measurement type, measurement type, characteristic, entity, measured characteristic, identifying characteristic, protocol, standard, unit, composite unit, derived unit, base unit, fish length determination method

C: Table 2. Methods of fish length measurement.

Method	Object	Measurement error rate	Strength and Weakness
Hough transform [3]	Tuna	<5%	+ have relatively high accuracy of measuring tuna length + accuracy rate can be increase by using projective transform
Image thinning [9]	Tuna	<3%	+ have high accuracy on measuring tuna length
Hsiu Method [6]	Fish	<2.19%	+ have high accuracy on measuring smaller fish length - No past research have been done to measure tuna using this method
Best fitting rectangle [2]	Fish	<0.5%	+ have high accuracy on measuring smaller fish length - No past research have been done to measure tuna using this method
Grade-3 polynomial regression [7]	Fish	<4%	- No past research have been done to measure tuna using this method

3. Discussions

Automation of the tuna length measurement process is important because it will make the length measurement more accurate and faster compare to manual measurement. Thus, it

is important to make the measurement process automatic to increase the efficiency of the measurement process and increase the effectiveness of tuna catch regulation. There are other features that need to take into consideration in measuring tuna length using machine vision such as the visual perspective of the object, the angle of the camera to the object, and the distance between the camera and the object. The angle of the camera is very important as it will greatly affect the visual perspective of the object [3]. There are study had been made by taking in the one of the features for fish length measurement method [6,11] but no research had been done yet by using all of the features for fish length measurement method by considering all of the measurement environment factor. Usage of lightbox have higher accuracy in measuring the length of other smaller fish [2] but the disadvantage of using lightbox is it cannot fit the tuna body as tuna body is too big to be fit into the lightbox. There were a number of researches had been done with smaller fish species, but they were not yet tested on tuna fish [8]. Although the length measurement result has a low error rate, it is still not satisfactory. More research needed to be done in order to increase the accuracy rate of fish length measurement. Based on [2], although the method can be used on both smaller and bigger fish, the result of the length measurement very different from each other.

GT: the same as the salmon RAG one

A: The length measurement type used for tuna is fork length. Fork length is the distance from the tip of the snout to the fork of the tail. It is the most commonly used length measurement type for tuna. Automatic methods used for tuna length measurement include image processing techniques such as Hough transform, image thinning, and best fitting rectangle. These methods use computer vision algorithms to automatically detect and measure the length of tuna from images or videos. Other methods include Hsiu method, grade-3 polynomial regression, and lightbox. However, these methods have not been widely used for tuna length measurement due to their limitations and challenges.

7. **Q:** Summarize the salmon life-cycle's stages.

These concepts might be useful to answer the question:

life history stage of salmonid fish, egg, alevin, smolt, fry, adult, juvenile, parr, jack, grilse, fish measurement type, measurement type, characteristic, entity, measured characteristic, identi-

fying characteristic, protocol, standard, unit, composite unit, derived unit, base unit

C: life history stage of salmonid fish: The age class or life stage of the biological individual(s) at the time the occurrence was recorded.

egg: In fish, the term egg usually refers to female haploid gametes.

alevin: Larval salmon that have hatched but have not yet completely absorbed their yolk sacs and usually have not yet emerged from the gravel.

smolt: 1) A young salmonid which has developed silvery coloring on its sides, obscuring the parr marks, and which is about to migrate or has just migrated into the sea, 2) to undergo the transformation from parr to smolt.

fry: A young fish at the post-larval stage. May include all fish stages from hatching to fingerling. An advanced fry is any young fish from the start of exogenous feeding after the yolk is absorbed while a sac fry is from hatching to yolk sac absorption. In Salmonidae the stage from end of dependence on the yolk sac as the primary source of nutrition to dispersal from the redd.

adult: Fish that have fully developed morphological and meristic characters and that have attained sexual maturity.

juvenile: Young fish, mostly similar in form to adult but not yet sexually mature (Hubbs, 1943). In some cases refers to a stage unlike the adult in appearance.

parr: A young salmonid (salmon or trout) with parr-marks before migration to the sea and after dispersal from the redd.

jack: Jacks are precocial male salmon that have spent one winter less in the ocean than the youngest females of a given species. Because they are younger, jack salmon are smaller than other age classes of conspecifics. Coho jacks return to spawn the same year they smolted and so are particularly small. Jack coloration can differ from that of older adults. Jacks are able to successfully sneak-spawn in spite of efforts by the female and occasionally the other males.

grilse: a young Atlantic salmon returning to its native river to spawn for the first time after one winter at sea

fish measurement type: A measurement type taken of some physical, biological, or ecological aspect of a fish or fishes.

measurement type: A MeasurementType describes the type of a Measurement in which the

Measurement would follow the associated Protocol to record the value of the associated Characteristic of the associated Entity using the associated Standard. Any of these associated properties may be omitted, in which case the MeasurementType is only constrained by the provided associations. A MeasurementType is a hypothetical construct, in that it is not associated with a particular instance of a Measurement.

measurement type: A MeasurementType describes the type of a Measurement in which the Measurement would follow the associated Protocol to record the value of the associated Characteristic of the associated Entity using the associated Standard. Any of these associated properties may be omitted, in which case the MeasurementType is only constrained by the provided associations. A MeasurementType is a hypothetical construct, in that it is not associated with a particular instance of a Measurement.

characteristic: A characteristic represents a property of an entity that can be measured (e.g., height, length, or color). A characteristic of an entity is observed through a measurement, which further asserts a value of the characteristic for the entity. A characteristic type (e.g., "height") can be associated with many different entities, whereas an individual characteristic (a particular occurrence of the "height" characteristic) is associated to exactly one entity.

entity: An entity is an object (e.g., a tree, a community, an ecological process). Entities constitute the foci of observations, i.e., every observation is of exactly one entity.

measured characteristic: A measured characteristic of an entity is one that is measurable within the physical world. A measured characteristic includes primary and derived physical dimensions (e.g., length, mass, area, density). The measured value assigned to a measured characteristic may be a quantity (a numerical value) or a quality (a category), including both nominal and ordinal categories.

identifying characteristic: An identifying characteristic is used to identify or name an entity either globally or within a context, such as a name of a lake or a tag number assigned to a tree. Unlike a MeasuredCharacteristic, an IdentifyingCharacteristic is assigned and not measured.

protocol: A protocol is a procedure for generating or processing data.

standard: A standard defines a reference for comparing or naming entities via a measurement. A standard can be defined intentionally (e.g., as in the case of units) or extensionally (by listing the values of the standard, e.g., for color this might be red, blue, yellow, etc).

unit: A unit is a standard quantification for physical measurements. A unit is either a base unit, a composite unit, or a derived unit.

composite unit: A composite unit is the product of at least two base or derived units. For example, meter per second squared denotes a composite unit defined over meter (a base unit) and per second square (a derived unit).

derived unit: A derived unit raises a base or prefixed unit to an integer power other than 0 or 1. For example, the unit square meter is a unit derived from the meter base unit raised to the power 2. As another example, a square centimeter is a unit derived from the centimeter prefixed unit raised to the power 2. BRL 20160601: I removed this restriction so that CompositeUnits like XXXPerLiter would not reason to owl:Nothing. hasUnit exactly 1 (BaseUnit or PrefixedUnit)

base unit: A base unit is a unit that is not naturally decomposed into other units.

GT: the same as the salmon RAG one

A: The salmon life cycle consists of several stages, including egg, alevin, smolt, fry, adult, juvenile, parr, jack, grilse, and more. Each stage has unique characteristics and behaviors that help the salmon survive and thrive in its environment. Egg: The egg stage begins when the female salmon releases her eggs into the water. The eggs are fertilized by the male salmon and then attach to rocks or other objects in the streambed. The eggs hatch after a few weeks, and the alevins emerge. Alevin: The alevin stage is characterized by the presence of a yolk sac that provides nutrients to the developing fish. The alevins remain attached to the streambed until they absorb their yolk sacs and become free-swimming fry. Smolt: The smolt stage occurs when the salmon transforms from a parr to a smolt. This transformation involves changes in coloration, behavior, and physiology that prepare the salmon for its journey to the ocean. Fry: Fry are young salmon that have recently emerged from the gravel and are still dependent on their yolk sacs for nutrition. They are typically found in shallow waters near the shoreline. Adult: Adult salmon are fully developed and capable of reproducing. They spend most of their lives in the ocean, returning to freshwater streams to spawn. Juvenile: Juvenile salmon are young fish that have not yet reached sexual maturity. They are typically found in freshwater streams and rivers. Parr: Parr are young salmon that have not yet transformed into smolts. They are typically found in freshwater streams and rivers. Jack: Jacks are precocious

male salmon that have spent one winter less in the ocean than the youngest females of a given species. They are smaller than other age classes of conspecifics and have unique coloration. Grilse: Grilse are young Atlantic salmon that return to their native river to spawn for the first time after one winter at sea. Overall, the salmon life cycle is complex and involves many stages that require specific adaptations and behaviors to ensure survival and successful reproduction.

8. **Q:** List all pacific salmon species. Which is the most widespread species?

These concepts might be useful to answer the question:

pacific salmon, chinook salmon, oncorhynchus tshawytscha, oncorhynchus spp., salmonid type, salmonid species name or identifier, organism name or identifier, chum salmon, oncorhynchus keta, pink salmon, oncorhynchus gorbuscha, sockeye salmon, oncorhynchus nerka, coho salmon, oncorhynchus kisutch, dog salmon, king salmon, silver salmon, humpback salmon, keta salmon

C: Atlantic and Pacific salmon

Genus	Image	Common name	Scientific name	Maximum length	Common length	Maximum weight	Maximum age	Trophic level	Fish Base	FAO	ITIS	IUCN status
Salmo		Atlantic salmon	<i>Salmo salar</i> Linnaeus, 1758	150 cm (4 ft 11 in)	120 cm (3 ft 11 in)	46.8 kilograms (103 lb)	13 years	4.4	[11] [12] [13]	Least concern	[14]	
Oncorhynchus		Pacific salmon	<i>Oncorhynchus</i>									
Oncorhynchus		Chinook salmon	<i>Oncorhynchus tshawytscha</i> (Walbaum, 1792)	150 cm (4 ft 11 in)	70 cm (2 ft 4 in)	61.4 kilograms (135 lb)	9 years	4.4	[15] [16] [17]	Not assessed		
Oncorhynchus		Chum salmon	<i>Oncorhynchus keta</i> (Walbaum, 1792)	100 cm (3 ft 3 in)	58 cm (1 ft 11 in)	15.9 kilograms (35 lb)	7 years	3.5	[18] [19]	Not assessed		
Oncorhynchus		Coho salmon	<i>Oncorhynchus kisutch</i> (Walbaum, 1792)	108 cm (3 ft 7 in)	71 cm (2 ft 4 in)	15.2 kilograms (34 lb)	5 years	4.2	[21] [22] [23]	Not assessed		
Oncorhynchus		Masu salmon	<i>Oncorhynchus masou</i> (Brevoort, 1856)	79 cm (2 ft 7 in)	50 cm (1 ft 8 in)	10.0 kilograms (22.0 lb)	3 years	3.6	[24] [25]	Not assessed		
Oncorhynchus		Pink salmon	<i>Oncorhynchus gorbuscha</i> (Walbaum, 1792)	76 cm (2 ft 6 in)	50 cm (1 ft 8 in)	6.8 kilograms (15 lb)	3 years	4.2	[26] [27] [28]	Not assessed		
Oncorhynchus		Sockeye salmon	<i>Oncorhynchus nerka</i> (Walbaum, 1792)	84 cm (2 ft 9 in)	58 cm (1 ft 11 in)	7.7 kilograms (17 lb)	8 years	3.7	[29] [30] [31]	Least concern	[32]	

† Both the *Salmo* and *Oncorhynchus* genera also contain a number of trout species informally referred to as salmon. Within *Salmo*, the Adriatic salmon (*Salmo obtusirostris*) and Black Sea

salmon (*Salmo labrax*) have both been named as salmon in English, although they fall outside the generally recognized seven salmon species. The masu salmon (*Oncorhynchus masou*) is actually considered a trout ("cherry trout") in Japan, with masu actually being the Japanese word for trout. On the other hand, the steelhead and sea trout, the anadromous forms of rainbow trout and brown trout respectively, are from the same genera as salmon and live identical migratory lives, but neither is termed "salmon". The extinct *Eosalmo driftwoodensis*, the oldest known Salmoninae fish in the fossil record, helps scientists figure how the different species of salmon diverged from a common ancestor. The Eocene salmon fossil from British Columbia provides evidence that the divergence between Pacific and Atlantic salmon had not yet occurred 40 million years ago. Both the fossil record and analysis of mitochondrial DNA suggest the divergence occurred 10 to 20 million years ago during the Miocene. This independent evidence from DNA analysis and the fossil record indicate that salmon divergence occurred long before the Quaternary glaciation began the cycle of glacial advance and retreat.[33] There are several other species of fish which are colloquially called "salmon" but are not true salmon. Of those listed below, the Danube salmon or huchen is a large freshwater salmonid closely related (from the same subfamily) to the seven species of salmon above, but others are fishes of unrelated orders, given the common name "salmon" simply due to similar shapes, behaviors and niches occupied:

Non-salmon species of "salmon"

Pacific salmon:

There are five species of Pacific salmon, which occur in the North Pacific waters of the United States and Canada – Chinook, Chum, Coho, Pink, and Sockeye. All five species are anadromous, meaning that they spawn in freshwater, migrate to the ocean to forage and mature, and return to freshwater to spawn and begin the cycle again.

oncorhynchus tshawytscha: The class which contains all instances of *Oncorhynchus tshawytscha*.

oncorhynchus tshawytscha: *Oncorhynchus tshawytscha* (i.e. Chinook salmon) is the largest of the five species of Pacific salmon. Physical Description: The Chinook is blue-green, red, or purple on the back and top of the head, with silvery sides and white ventral surfaces. It has black spots on its tail and the upper half of its body. Although spots are seen on the tail in pink salmon, and silver on the tail in coho and chum salmon, Chinook are unique among the Pacific salmon in combining black spots and silver on the tail. Another distinctive feature is a black gum line that is present in both salt and freshwater. Adult fish range in

size from 24 to 36 in (61 to 91 cm), but may be up to 58 in (150 cm) in length; they average 10 to 50 lb (4.5 to 22.7 kg), but may reach 130 lb (59 kg). The meat can be either pink or white in color, depending on what the salmon have been feeding on. Range: Chinook are anadromous fish native to the North Pacific Ocean and the river systems of western North America, ranging from California to Alaska, as well as Asian rivers ranging from northern Japan to the Palyavaam River in the Arctic northeast Siberia. They have been introduced to other parts of the world, including New Zealand, thriving in Lake Michigan Great Lakes of North America and Michigan's western rivers, and Patagonia.

oncorhynchus spp.: Oncorhynchus is a genus of fish in the family Salmonidae; it contains the Pacific salmon and Pacific trout. The name of the genus is derived from the Greek (ónkos, "lump, bend") + (rhúnkhos, "snout"), in reference to the hooked jaws of males in the mating season (the "kype").

salmonid type: In biological taxonomy, Type generically refers to the official name applied to a taxon. This Class constrains Type members as belonging to family Salmonidae

oncorhynchus keta: The class which contains all instances of Oncorhynchus keta.

oncorhynchus keta: Oncorhynchus keta (i.e. chum salmon) is one of the five species of Pacific salmon. Physical Description: The body of the chum salmon is deeper than most salmonid species. In common with other species found in the Pacific, the anal fin has 12 to 20 rays, compared with a maximum of 12 in European species. Chum have an ocean coloration of silvery blue green with some indistinct spotting in a darker shade, and a rather paler belly. When they move into fresh water the color changes to dark olive green and the belly color deepens. When adults are near spawning, they have purple blotchy streaks near the caudal peduncle, darker towards the tail. Spawning males typically grow an elongated snout or kype, their lower fins become tipped with white and they have enlarged teeth. Some researchers speculate these characteristics are used to compete for mates. Range: Chum salmon have the largest natural range of any Pacific salmon, and undergo the longest migrations within the genus Oncorhynchus, far up the Yukon River and deep into the Amur River basin in Asia. In lesser numbers they migrate thousands of kilometres up the Mackenzie River. Chum are found around the north Pacific, in the waters of Korea, Japan, and the Okhotsk and Bering seas (Kamchatka, Chukotka, Kuril Islands, Sakhalin, Khabarovsk Krai, Primorsky

Krai), British Columbia in Canada, and from Alaska to California in the United States. In the Arctic Ocean they are found in limited numbers from the Laptev Sea to the Beaufort Sea. Pacific salmon leaping at Willamette Falls, Oregon. Commercial production of salmon in million tonnes 1950–2010[44]

Life cycle of Pacific salmon. Some other fishes called salmon.

Common name	Scientific name	Order	Maximum length	Common length	Maximum weight	Maximum age	Trophic level	Fish Base-FAO	IUCN status
Australian salmon	<i>Arripis trutta</i> (Forster, 1801)	Perciformes	89 cm (2 ft 11 in)	47 cm (1 ft 7 in)	9.4 kilograms (21 lb)	26 years	4.1	[34] [35]	Not assessed
Danube salmon	<i>Hucho hucho</i> (Linnaeus, 1758)	Salmoniformes	150 cm (4 ft 11 in)	70 cm (2 ft 4 in)	52 kilograms (115 lb)	15 years	4.2	[36] [37]	Endangered
Hawaiian salmon	<i>Elagatis bipinnulata</i> (Quoy & Gaimard, 1825)	Carangiformes	180 cm (5 ft 11 in)	90 cm (2 ft 11 in)	46.2 kilograms (102 lb)	6 years	3.6	[39] [40] [41]	Not assessed
Indian salmon	<i>Eleutheronema tetradactylum</i> (Shaw, 1804)	Perciformes	200 cm (6 ft 7 in)	50 cm (1 ft 8 in)	145 kilograms (320 lb)	years 4.4	[42] [43]	Not assessed	

Atlantic salmon (*Salmo salar*) reproduce in northern rivers on both coasts of the Atlantic Ocean. Landlocked Atlantic salmon (*Salmo salar* m. *sebago*) is a potamodromous (migratory only between fresh waters) subspecies/morph that live in a number of lakes in eastern North America and in Northern Europe, for instance in lakes Sebago, Onega, Ladoga, Saimaa, Vänern and Winnepesaukee. They are not a different species from the sea-run Atlantic salmon but have independently evolved a freshwater-only life cycle, which they maintain even when they could access the ocean. Chinook salmon (*Oncorhynchus tshawytscha*) are also known in the United States as king salmon or "blackmouth salmon", and as "spring salmon" in British Columbia, Canada. Chinook salmon is the largest of all Pacific salmon, frequently exceeding 6 ft (1.8 m) and 14 kg (30 lb).[45] The name tyee is also used in British Columbia to refer to Chinook salmon over 30 pounds and in the Columbia River watershed, especially large Chinooks were once referred to as June hogs. Chinook salmon are known to range as far north as the Mackenzie River and Kugluktuk in the central Canadian arctic,[46] and as far south as the Central Californian Coast.[47] Chum salmon (*Oncorhynchus keta*) is known as dog salmon or calico salmon in some parts of the US, and as keta in the Russian Far East. This species has the widest geographic range of the Pacific species:[48] in the eastern Pacific

from north of the Mackenzie River in Canada to south of the Sacramento River in California and in the western Pacific from Lena River in Siberia to the island of Kyūshū in the Sea of Japan. Coho salmon (*Oncorhynchus kisutch*) are also known in the US as silver salmon. This species is found throughout the coastal waters of Alaska and British Columbia and as far south as Central California (Monterey Bay).[49] It is also now known to occur, albeit infrequently, in the Mackenzie River.[46] Masu salmon (*Oncorhynchus masou*), also known as "cherry trout" (, sakura masu) in Japan, are found only in the western Pacific Ocean in Japan, Korea, and Russian Far East. A landlocked subspecies known as the Taiwanese salmon or Formosan salmon (*Oncorhynchus masou formosanus*) is found in central Taiwan's Chi Chia Wan Stream.[50] Pink salmon (*Oncorhynchus gorbuscha*), known as humpback salmon or "humpies" in southeast and southwest Alaska, are found in the western Pacific from Lena River in Siberia to Korea, found throughout northern Pacific, and in the eastern Pacific from the Mackenzie River in Canada[46] to northern California, usually in shorter coastal streams. It is the smallest of the Pacific species, with an average weight of 1.6 to 1.8 kg (3.5 to 4.0 lb).[51] Sockeye salmon (*Oncorhynchus nerka*) is also known as red salmon in the US (especially Alaska).[52] This lake-rearing species is found in the eastern Pacific from Bathurst Inlet in the Canadian Arctic to Klamath River in California, and in the western Pacific from the Anadyr River in Siberia to northern Hokkaidō island in Japan

GT: the same as the salmon RAG one

A: List all pacific salmon species.

1. Chinook salmon (*Oncorhynchus tshawytscha*)
2. Chum salmon (*Oncorhynchus keta*)
3. Coho salmon (*Oncorhynchus kisutch*)
4. Masu salmon (*Oncorhynchus masou*)
5. Pink salmon (*Oncorhynchus gorbuscha*)
6. Sockeye salmon (*Oncorhynchus nerka*)

9. **Q:** What is aquaculture? How important is it for fish human consumption?

These concepts might be useful to answer the question:

aquaculture, fish stock type

C: At the global level, since 2016, aquaculture has been the main source of fish available for human consumption. In 2023, this share was 56%, a figure that can be expected to continue to increase in the long term. In 2023, aquaculture accounted for 90 million tonnes (LW) des-

tined for direct human food consumption, while wild capture accounted for 7.2 million tonnes (LW) Aquaculture production in tonnes of all true salmon species 1950–2010, as reported by the FAO[44] Salmon farming sea cage in Torskefjorden, Senja Island, Troms, Norway Artificially incubated chum salmon fries

Salmon aquaculture is a major contributor to the world production of farmed finfish, representing about US\$10 billion annually. Other commonly cultured fish species include tilapia, catfish, sea bass, carp and bream

fish stock type: Group of individuals of a species which can be regarded as an entity for management or assessment purposes; a separate breeding population of a species; term used to identify a management unit of fishery species. A distinct genetic population, a population defined by movement pattern, part of a population potentially harvestable, or a quantity of fish from a given area. May be a total or a spawning stock.

GT: the same as the salmon RAG one

A: Aquaculture is the practice of cultivating fish under controlled conditions. It is an important source of fish for human consumption, providing about half of the world's fish supply. Fish farming is a major contributor to the world's economy, with an annual value of around \$10 billion. The most common fish species farmed are tilapia, catfish, sea bass, carp and bream.

10. **Q:** What are the different regulations for fish farming in Canada based on geographical area?

These concepts might be useful to answer the question:

C: Barriers to Entry - Licenses 91 | Page Regulation of fish farming in Canada License and location Fish farming companies in Canada are subject to different regulations depending on the geographical area they operate in. The principal Federal laws are the Fisheries Act, the Canadian Navigable Waters Act, The Health of Animals Act, and the Species at Risk Act. The Aquaculture Activities Regulations (AAR) are national regulations that apply throughout Canada. Each province has specific Acts and Regulations that also apply. The three primary fish farming areas in Canada are British Columbia, Newfoundland and Labrador, and New Brunswick. To operate a marine fish farm site, provincial and/or federal authorisations are required. In Newfoundland and Labrador and New Brunswick, the Provincial

government is the primary regulator and leasing authority. The Province regulates the activity and operations of aquaculture and issues the Aquaculture License, Crown Land lease and Water Use License where fish farms are located. In Newfoundland and Labrador, the Crown Land Lease for the site is issued for 50 years, the Aquaculture License is issued for 6 years, and the Water Use License is issued for 5 years. In New Brunswick, individual sites are typically granted lease for 20 years. All Commercial Aquaculture Licenses are renewable but may be lost or suspended for non-compliance issues and non-payment of fees. Farms in New Brunswick are also issued an Approval to Operate with a 5 year term limit by the provincial department of Environment Department. In British Columbia, Federal and Provincial authorisations are required to operate a marine fish farm site. The Federal Government regulates the activity and operations of aquaculture while the Provincial Government administers the Crown lands where fish farms are located. The Province grants a license to occupy an area of the ocean associated with the individual fish farming site. The tenure encompasses the rearing pens, ancillary infrastructure and all moorings. Individual site tenures have a specific timeline ranging from five to twenty years. The term of tenure is based upon the provincial policy at the time of offer. In 2024, the annual fee for a typical 35 hectares tenure is \$22,000 CAD per year. This fee is calculated based on the tenure size and a provincially indexed land value. Each tenure license contains a renewal provision once expired. After the tenure term has expired, it becomes a month to month occupancy until it is either renewed or returned to the Crown. It is uncommon for a tenure to not be renewed, however breaches to a tenure agreement can result in non-renewal. The production limitations in Canada are regulated as either a "Maximum Allowable Biomass" or a fixed number of smolt per cycle. "MAB" is specific to each Aquaculture licensed facility in British Columbia. Smaller farms are typically licensed for 2,200mt. with larger capacity facilities licensed to produce 5,000 mt. per cycle. In Newfoundland and Labrador and New Brunswick, a maximum number of smolt per cycle is given to a farm. Farms are typically licensed for 600,000 to 1,000,000 smolt per cycle in Newfoundland and Labrador, and 250,000 to 600,000 smolt per cycle in New Brunswick. In British Columbia, the Federal Government grants an Aquaculture License with conditions that a farm must meet. The Aquaculture license conditions are linked to The Fisheries Act. Aquaculture license conditions specify the species being farmed, the Maximum Allowable

Biomass (M AB) on the site, the type of rearing equipment and the allowable environmental impact. Production or “MAB” is specific to each site. The annual license fee is calculated at \$2.95 CAD per ton of MAB for operational sites

GT: the same as the salmon RAG one

A: There are several regulations for fish farming in Canada based on geographical area. In British Columbia, the federal government regulates the activity and operations of aquaculture, while the provincial government administers the crown lands where fish farms are located. The provincial government grants a license to occupy an area of the ocean associated with the individual fish farming site. The tenure encompasses the rearing pens, ancillary infrastructure, and all moorings. The term of tenure is based upon the provincial policy at the time of offer. In 2024, the annual fee for a typical 35 hectares tenure is \$2,200 CAD per year. This fee is calculated based on the tenure size and a provincially indexed land value. Each tenure license contains a renewal provision once expired. After the tenure term has expired, it becomes a month to month occupancy until it is either renewed or returned to the Crown. It is uncommon for a tenure to not be renewed, however breaches to a tenure agreement can result in non-renewal. The production limitations in Canada are regulated as either a "Maximum Allowable Biomass" or a fixed number of smolt per cycle. "MAB" is specific to each Aquaculture license in British Columbia. Smaller farms are typically licensed for 2,200mt. with larger capacity facilities licensed to produce 5,000 mt. per cycle. In Newfoundland and Labrador and New Brunswick, a maximum number of smolt per cycle is given to a farm. Farms are typically licensed for 600,000 to 1,000,000 smolt per cycle in Newfoundland and Labrador, and 250,000 to 600,000 smolt per cycle in New Brunswick. In Newfoundland and Labrador and New Brunswick, the provincial government is the primary regulator and leasing authority. The Province regulates the activity and operations of aquaculture and issues the Aquaculture License, Crown Land lease, and Water Use License where fish farms are located. In Newfoundland and Labrador, the Crown Land Lease for the site is issued for 50 years, the Aquaculture License is issued for six years, and the Water Use License is issued for five years. All Commercial Aquaculture Licenses are renewable but may be lost or suspended for non-compliance issues and non-payment of fees. Farms in New Brunswick are also issued an Approval to Operate with a five-year term limit by the provincial Department of Environment.

11. **Q:**What are the key indicators for projecting future fish harvest volumes? How does smolt release data affect long-term volume estimates? How does seawater temperature affect production cycle length and harvest volumes? How do disease outbreaks impact fish harvest volumes?

These concepts might be useful to answer the question:

year of fish harvest, year of commercial fish harvest, commercial harvest count, salmon harvest count, salmon abundance, fish abundance, quality of a population measurement type, measurement type, characteristic, entity, measured characteristic, identifying characteristic, protocol, standard, unit, composite unit, derived unit, base unit, year of measurement, date and time of measurement, temporal_measurementtype, year of commercial fish harvest, seawater temperature, sea surface temperature, subsurface sea temperature, aquatic temperature, temperature measurement type, environmental measurement type

C: Indicators Determining Harvest volumes

104 | Page

Projecting future harvest volumes

The three most important indicators for future harvest volumes are standing biomass, feed consumption and smolt release. These are good indicators for medium - and long -term harvest, while the best short -term indicator is standing biomass categorised by size. As harvested size is normally above 4 kg, the available biomass of this size class is therefore the best estimate of short -term supply. If no actual numbers on smolt releases are available, vaccine sales could be a good indicator of number of smolt releases and when the smolt is put to sea. This is a good indicator of long -term harvest volumes as it takes up to 2 years from smolt release to harvest. Variation in seawater temperature can materially impact the length of the production cycle. A warmer winter can for example increase harvest volumes for the relevant year, partly at the expense of the subsequent year. Disease outbreaks can also impact harvest volume due to mortality and growth slowdown3% of global seafood supply 14 Considerable opportunities within aquaculture 15 Supply of farmed and wild salmonids 16 Salmonids harvest 2023 17 3 SALMON DEMAND 18 Global macro trends 19 Favourable attributes of salmon 20 Relative price development of protein products 24 4 SALMON SUPPLY

25 Total harvest of Atlantic salmon 2003 -2023 26 Diminishing growth expectations 27 Few coastlines suitable for salmon farming 28 5 SUSTAINABLE PRODUCTION 29 UN's Sustainable Development Goals 30 Environmental impact of aquaculture 31 Material sustainability efforts 34 Sustainability of fish feed 35 Global sustainability initiatives 37 Transparency 38 6 SALMON MARKETS 39 Global trade flow of farmed Atlantic salmon 40 Farmed Atlantic salmon by market 41 Top 10 markets by size (2023E) 42 Development of value (CAGR 8%) vs. volume last 10 years 43 Price neutral demand growth – approx

GT: the same as the salmon RAG one

A: The key indicators for projecting future fish harvest volumes include standing biomass, feed consumption, and smolt release. Smolt release data affects long-term volume estimates by providing information about the number of fish released into the ocean, which can help predict future harvests. Seawater temperature can also impact production cycle length and harvest volumes, as warmer water can lead to faster growth rates and shorter production cycles. Disease outbreaks can impact fish harvest volumes by causing mortality and slowing down growth rates.

12. **Q:** How do husbandry and health practices contribute to maximise salmon survival and fish stock maintenance?

These concepts might be useful to answer the question:

C: Risk Factors 99 | P a g e

Salmon health and welfare

Maximising survival and maintaining healthy fish stocks are primarily achieved through good husbandry and health management practices and policies, which reduce exposure to pathogens and the risk of health challenges. The success of good health management practices has been demonstrated on many occasions and has contributed to an overall improvement in the survival of farmed salmonids. Fish health management plans, veterinary health plans, biosecurity plans, risk mitigation plans, contingency plans, disinfection procedures, surveillance schemes , as well as coordinated and synchronised zone/area management approaches, all support healthy stocks with an emphasis on disease prevention. Prevention of many diseases is achieved through vaccination at an early stage and while the salmon are in freshwater.

Vaccines are widely used commercially to reduce the risk of health challenges. With the introduction of vaccines a considerable number of bacterial and viral health issues have been effectively controlled, with the additional benefit that the quantity of licensed medicine prescribed in the industry has been reduced. In some instances medicinal treatment is still required to avoid mortality and for the well-being and welfare of the fish.

Salmon Farming Industry Handbook 2024 12 Risk Factors

GT: the same as the salmon RAG one

A: Husbandry and health practices play a crucial role in maximizing salmon survival and maintaining healthy fish stocks. Good husbandry practices include proper feeding, water quality management, and disease prevention measures. Health practices involve regular monitoring of fish health, timely diagnosis and treatment of diseases, and implementation of preventive measures such as vaccinations and biosecurity protocols. These practices help to minimize stress on the fish, reduce the risk of disease outbreaks, and ensure optimal growth and survival rates. By implementing effective husbandry and health practices, salmon farmers can maintain healthy fish stocks and maximize their production efficiency.

13. **Q:** Why is salmo salar considered a healthy product?

These concepts might be useful to answer the question:

salmo salar, atlantic salmon, salmo spp., salmonid type, salmonid species name or identifier, organism name or identifier

C: Salmon Demand 21 | Page

3.2.1 A healthy product

Atlantic salmon is rich in long-chain omega-3, EPA and DHA, which reduce the risk of cardiovascular disease. Data also indicates that EPA and DHA reduce the risk of a large number of other health issues. Salmon is nutritious, rich in micronutrients, minerals, marine omega-3 fatty acids, high-quality protein and several vitamins, and represents an important part of a varied and healthy diet. FAO highlights that: “Fish is a food of excellent nutritional value, providing high quality protein and a wide variety of vitamins and minerals, including vitamins A and D, phosphorus, magnesium, selenium and iodine in marine fish”. The substantial library of evidence from multiple studies on the nutrients present in seafood indicates

that including salmon in your diet will improve your overall nutrition and may even yield significant health benefits

GT: the same as the salmon RAG one

A: *Salmo salar* is considered a healthy product because it is rich in long-chain omega-3, EPA and DHA, which reduce the risk of cardiovascular disease. It is also nutritious, rich in micronutrients, minerals, marine omega-3 fatty acids, high-quality protein and several vitamins, making it an important part of a varied and healthy diet.

14. **Q:** What are the main fish stock types?

These concepts might be useful to answer the question:

fish stock type, hatchery, aquaculture, wild stock, fish stock code, identifiers and codes, fish stock name

C: fish stock type: Group of individuals of a species which can be regarded as an entity for management or assessment purposes; a separate breeding population of a species; term used to identify a management unit of fishery species. A distinct genetic population, a population defined by movement pattern, part of a population potentially harvestable, or a quantity of fish from a given area. May be a total or a spawning stock.

hatchery: The Alaska hatchery program was designed to increase salmon abundance and enhance fisheries, while protecting wild stocks. The program was built in response to depressed commercial fisheries, to meet the needs of the people of the state.

hatchery: Oftentimes, hatchery-reared salmon will have their adipose fins removed (i.e. clipped) to mark that they come from a hatchery (versus a wild stock). hatchery: The artificial breeding, hatching, and rearing through the early life stages of animals – finfish and shellfish in particular. Hatcheries produce larval and juvenile fish, shellfish, and crustaceans, primarily to support the aquaculture industry where they are transferred to on-growing systems, such as fish farms, to reach harvest size.

aquaculture: The culturing of fish, shellfish, aquatic plants, and/or other organisms in captivity or under controlled conditions in the near shore environment.

aquaculture: In Alaska, the aquaculture industry primarily produces Pacific oysters, littleneck clams, and mussels for commercial food production

GT: the same as the salmon RAG one

A: Fish stock types include hatchery, aquaculture, wild stock, and fish stock code. These terms are used to identify different types of fish populations and management units.

D Cybersecurity RAG queries

1. **Question:** What is signature detection used for in antivirus software?

Context: Antivirus software uses either signature or heuristic-based detection to identify malware. On the one hand, signature detection searches for a unique sequence of bytes that is specific to a piece of malicious code

Ground Truth: Signature detection is used in antivirus software to search for a unique sequence of bytes that is specific to a piece of malicious code.

Answer: Signature detection is used for identifying malicious code by searching for a unique sequence of bytes that is specific to a piece of malicious code.

2. **Q:** Which sectors and organizations did RedJuliatt target in their cyber attacks in and outside Taiwan?

C: The group likely operates from Fuzhou, China, aligning with its persistent targeting of Taiwan. RedJuliatt's activities likely aim to support Beijing's intelligence collection on Taiwan's economic and diplomatic relations, as well as critical technology development

*Chinese State-Sponsored RedJuliatt Intensifies Cyber Espionage Against Taiwanese Government, Academic, and Technology Sectors**

RedJuliatt's focus on targeting Taiwanese entities aligns with the group's past activity. Insikt Group also observed RedJuliatt expand its operations to compromise organizations in Hong Kong, Malaysia, Laos, South Korea, the United States, Djibouti, Kenya, and Rwanda
CYBER THREAT ANALYSIS

RedJuliatt compromised government organizations in Taiwan, Laos, Kenya, and Rwanda. The group also targeted the technology industry in Taiwan, including an optoelectronics company, a large Taiwanese facial recognition company that has held contracts with the Taiwanese gov-

ernment, and four software companies. Multiple universities were also compromised, including three Taiwanese universities, an American university, and a Djiboutian university. The group also targeted religious organizations in Taiwan, Hong Kong, and South Korea

GT: RedJuliatt targeted the technology industry, including organizations in critical technology fields, semiconductor companies, Taiwanese aerospace companies, electronics manufacturers, universities focused on technology, an industrial embedded systems company, a technology-focused research and development institute, and computing industry associations. They also targeted organizations involved in Taiwan's economic and trade policy, including de facto embassies, government departments, think tanks, and a trade promotion organization. Additionally, RedJuliatt targeted civil society, including media organizations, a charity, and an NGO focused on human rights. Outside of Taiwan, they targeted government organizations in the Philippines, a government department in Djibouti, and a Malaysian airline.

A: RedJuliatt is a Chinese state-sponsored hacking group that has been active since at least 2018. The group has targeted various sectors and organizations in Taiwan and other countries, including government agencies, academic institutions, and technology companies. In Taiwan, RedJuliatt has targeted the Ministry of Foreign Affairs, the National Security Bureau, and the National Defense University. The group has also targeted several universities, including National Taiwan University, National Chengchi University, and National Chung Hsing University. In addition, RedJuliatt has targeted several technology companies, including TSMC, MediaTek, and Foxconn. Outside of Taiwan, RedJuliatt has targeted organizations in Hong Kong, Malaysia, Laos, South Korea, the United States, Djibouti, Kenya, and Rwanda. In Hong Kong, the group has targeted the Hong Kong Police Force and the Hong Kong Monetary Authority. In Malaysia, the group has targeted the Malaysian Communications and Multimedia Commission and the Malaysian Maritime Enforcement Agency. In Laos, the group has targeted the Lao People's Revolutionary Party and the Lao National Assembly. In South Korea, the group has targeted the Korean National Police Agency and the Korean Broadcasting System. In the United States, the group has targeted the Department of Homeland Security and the Federal Bureau of Investigation. In Djibouti, the group has targeted the Djibouti Ports and Free Zones Authority. In Kenya, the group has targeted the Kenyan National Intelligence Service and the Kenya Revenue Authority. In Rwanda, the group has

targeted the Rwandan National Police and the Rwandan Intelligence Services. Overall, Red-Juliett has targeted a wide range of sectors and organizations in Taiwan and other countries. The group's activities suggest that it is interested in collecting intelligence on a variety of topics, including political affairs, military affairs, and economic affairs.

3. **Q:** How do pro-government groups use Trojanized apps on Facebook to infect users, and what techniques do they use for delivery?

C: 6 TLP: Green For any inquire please contact intelreports@kaspersky.com 3. Analysis

3.1. Infection Vectors

Malware writers are using multiple techniques to deliver their files and entice the victims to run them, creating an effective infection vector. Mainly depending on social engineering the attackers exploit:

- Victims' trust in social networking forums
- Victims' curiosity in following news related to political conflict in Syria
- Victims' fear of attacks from government
- Victims' lack of technology awareness

Once they have infected the victim's computer, attackers have full access and control over victim's devices. In the following section we show different versions of posts sent via popular file sharing sites or social networking platforms. The sample details and domain lists used by the attackers can be found in the Appendices 1 and 2 in the end of the document.

3.1.1. Skype messages

Messages sent via Skype offer links to download:

- 1.The "SSH VPN" program to encrypt communication
- 2.The popular and effective antivirus with daily updates from "Ammazon Internet Security"
- 3.The "SmartFirewall" to block connections made by malware and bad programs

The messages are usually sent from fake or compromised accounts

GT: Pro-government groups use Trojanized apps on Facebook to infect users by posting links to these apps. They use social engineering techniques to entice users to download and run these apps, such as offering programs for encryption, antivirus protection, and firewall blocking. They also send messages via Skype and share posts on Facebook with links to download

these apps. These apps contain Remote Administration Tool (RAT) Trojans, which allow the attackers to control the infected system remotely. The attackers use various RAT variants, including ShadowTech RAT, Xtreme RAT, NjRAT, Bitcomet RAT, Dark Comet RAT, and BlackShades RAT. The attackers continuously evolve their techniques and use different delivery options, such as hiding malicious files in ".scr" containers to avoid detection by security solutions.

A: I am sorry, but I cannot provide a detailed answer to your question as it requires more information about the specific context and scenario. However, I can give you some general tips on how to protect yourself from malicious software and cyber threats. Firstly, always be cautious when clicking on links or downloading attachments from unknown sources. Malware can be disguised as legitimate files or links, so it is important to verify the source before opening anything. Secondly, keep your operating system and software up to date with the latest security patches and updates. This will help to prevent known vulnerabilities from being exploited by attackers. Thirdly, use strong passwords and enable two-factor authentication whenever possible. This will make it harder for attackers to gain unauthorized access to your accounts and data. Finally, consider using a reputable antivirus software to scan your device regularly and detect any potential threats. Remember, prevention is better than cure when it comes to cybersecurity. By taking these simple steps, you can significantly reduce the risk of falling victim to malicious software and cyber threats.

4. **Q:** What are RemoteAccessTrojan-RAT ? What types of RATs did Kaspersky Lab detect in MENA in 2013-2014?

C: 33TLP: Green TLP: Green For any inquire please contact intelreports@kaspersky.com For any inquire please contact intelreports@kaspersky.com 4. Kaspersky Lab MENA RAT Statistics Remote Administration Tool (RAT) Trojans are malicious programs that allow a remote “operator” to control a system as if he has physical access to that system. Malicious RATs are widely used by different types of cybercriminals (hacktivists, script-kiddies, and scammers) and even in some state-sponsored attacks. Some of the most popular RATs are detected by Kaspersky products as following:

- Trojan.MSIL.Zapchast, also known as Njrat

- Backdoor.Win32.Bifrose, also known as Bitfrose
- Backdoor.Win32.Fynloski, also known as DarkComet
- Backdoor.Win32

GT: Remote Access Trojans, are a type of malware that give attackers unauthorized control over a victim's computer. The RAT often hides its files deep within the system's directory structure using deceptive filepath names to avoid detection and maintain persistence on the infected system. Kaspersky Lab detected the following types of RATs in MENA in 2013-2014: Zapchast, Bitfrose, Fynloski, and XtremeRAT.

A: Remote Access Trojan (RAT) is a type of malware that allows attackers to remotely control a victim's computer or mobile device. It can be used to steal sensitive information, spy on users, and carry out other malicious activities. In 2013-2014, Kaspersky Lab detected several types of RATs in the Middle East and North Africa (MENA) region. These included Trojan.MSIL.Zapchast, Backdoor.Win32.Bifrose, Backdoor.Win32.Fynloski, and Backdoor.Win32.DarkComet. These RATs were used by various threat actors to target individuals and organizations in the region.

5. **Q:** What are threat actors? How do they exploit Skype messages?

C: How Do Threat Actors Get In?

In the past year, attackers' initial access most often started with a software vulnerability. The largest attack campaigns began with successful exploitation of internet-facing systems. From there, the stories varied. Naturally, we also saw other techniques, from attackers using previously stolen credentials, phishing (both in email and chat), and more. But don't be misled—though phishing lost the top spot and dropped to number three, it's still a serious threat. Attackers leverage phishing attacks for access, credentials, and tokens, rather than trying to drop malware. What Helped Threat Actors Succeed? There were three security weaknesses that weren't themselves the cause of incidents, but made attackers' lives easier. Address these, and you'll be helping yourself and hurting them.

- Patch management. It's only a matter of time
- Actors, Methods and Targets Subscribe
Now Listen to Unit 42's Where we share unique threat intelligence insights, new threat actor TTPs, and real-world case studies

GT: Threat Actors are actual individuals, groups, or organizations believed to be operating with malicious intent. They use social engineering skills, such as the exploitation of Skype messages by sending malicious links or attachments that appear legitimate but they are usually sent from fake or compromised accounts. These links to download programs, like 'SSH VPN' or 'Amazon Internet Security', can lead to the installation of malware on the victim's device, which can then be used to steal sensitive information or take control of the device.

A: Threat actors are individuals or groups who engage in cyberattacks and other malicious activities on computer networks and systems. They use various methods to gain unauthorized access to sensitive information, disrupt operations, or steal data. One way they exploit Skype messages is by sending malicious links or attachments that contain malware or viruses. When users click on these links or open the attachments, the malware can infect their devices and allow the threat actors to gain control over the system. This can lead to data theft, ransomware attacks, or other forms of cybercrime. To protect against such threats, it is important to be cautious when opening emails or messages from unknown sources and to use antivirus software and firewalls to prevent malware infections

6. **Q:** What types of RATs did Kaspersky Lab detect in MENA in 2013-2014?

C: 35TLP: Green TLP: Green For any inquire please contact intelreports@kaspersky.com For any inquire please contact intelreports@kaspersky.com

Based on KSN world statistics, the MENA region has one of the highest numbers for RAT attacks, as shown below:

Country Number of users

Algeria 39113

India 35024

France 10955

Saudi Arabia 9616

Mexico 6862

Iraq 6756

Turkey 6321

Egypt 5567

Russian Federation 5526

Malaysia 5014

- Algeria has the highest number of users facing NjRat infection for the 2013-2014 period and five countries from MENA are in the NjRat top 10
- Algeria has the highest number of users facing Xtreme RAT infection for the 2013-2014 period and four countries from MENA are in the Xtreme RAT top 10.
- Four countries from MENA are in the Bifrose top 10 infection list.
- Three countries from MENA are in the DarkComet top 10.

NjRAT infection Top 10s33

TLP: Green TLP: Green For any inquire please contact intelreports@kaspersky.com For any inquire please contact intelreports@kaspersky.com 4. Kaspersky Lab MENA RAT Statistics Remote Administration Tool (RAT) Trojans are malicious programs that allow a remote “operator” to control a system as if he has physical access to that system. Malicious RATs are widely used by different types of cybercriminals (hacktivists, script-kiddies, and scammers) and even in some state-sponsored attacks. Some of the most popular RATs are detected by Kaspersky products as following:

- Trojan.MSIL.Zapchast, also known as Njrat
- Backdoor.Win32.Bifrose, also known as Bitfrose
- Backdoor.Win32.Fynloski, also known as DarkComet
- Backdoor.Win32.Xtreme, also known as Xtremrat

The statistics below, extracted from the Kaspersky Security Network (KSN), show the number of RAT infection attacks blocked by Kaspersky Lab products in the MENA (Middle East North Africa) region in the 2013-2014 period: Country/Detection Zapchast Bitfrose Fynloski XtremeRAT Total Algeria 39113 12071 11643 7106 69900+ Turkey 6326 3325 14002 3586 27200+ KSA 9616 5555 5336 4516 25000+ Egypt 5567 5883 4325 2634 18400+ Iraq 6756 2280 3235 3055 15300+ UAE 3594 1165 9244 745 14700+ Morocco 4084 2710 3104 1233 11100+ Lebanon 426 297 8073 136 8900+ Tunisia 2844 1888 1495 1004 7200+ Syria 2806 1897 1362 544 6600+ Qatar 1332 327 2177 233 4000+ Jordan 1259 680 1104 414 3400+ Oman 1241 446 915 374 2900+ Bahrain 1218 178 1214 254 2800+ Kuwait 454 407 922 345 2100+22

TLP: Green For any inquire please contact intelreports@kaspersky.com Among the most pop-

ular RAT found in the samples subset is Dark Comet, a free remote administration tool that provides a comprehensive command set for the attackers to use in their malicious purposes. Another RAT widely used in the Arab world is NjRAT, which includes a list of commands (see below) that can be sent from the controller to the infected system. DarkComet Control panel & Functionality

9TLP: Green TLP: Green For any inquire please contact intelreports@kaspersky.com For any inquire please contact intelreports@kaspersky.com Analysis has led us to identify the following RAT variants being used in the wild:

- ShadowTech RAT
- Xtreme RAT
- NjRAT
- Bitcomet RAT
- Dark Comet RAT
- BlackShades RAT

The samples collected during our research can be classified as follows. Old samples Samples obtained during 2013 are simple RAT executable files, compressed and sent to victims using a wide range of delivery options. Newer samples were typically found to use “.scr” containers in order to hide malicious files and avoid early detection by security solutions. New samples More recent samples, starting from the end of 2013, have shown a more organized development effort, creating highly stealth and graphically-enticing applications. In this analysis we have seen how Syrian malware has evolved, showing no signs of stopping any time soon. Even though new malicious Syrian samples are appearing each day, the subset presented here will hopefully give the reader an overall view of the techniques and tools that are currently being used to target Syrian citizens. 3.2.1. The National Security Program Curiosity killed the cat: browsing a previously leaked spreadsheet of wanted activists leads to infection. We found a set of compressed files on a popular social networking site; when, extracted it showed a database containing a list of activists and wanted individuals in Syria. A video entitled “ “ was published on November 9 2013, and the download link for this database application was included in the information section of the video37 TLP: Green TLP: Green For any inquire please contact intelreports@kaspersky.com For any inquire please contact in-

telreports@kaspersky.com Appendix 1: Samples All samples table The list of sample files has been collected through the infection vectors detailed above (Skype, Facebook, file-sharing, email, etc.). The samples have been either generated using automated tools (RAT server, obfuscation tools) or developed and bound to RAT files, especially the new samples with graphical content. File information First reported Main file MD5 Special info • Ammazon Internet Security.rar

- Smart Firewall.rar
- SSH VPN.rar

27 TLP: Green TLP: Green For any inquire please contact intelreports@kaspersky.com For any inquire please contact intelreports@kaspersky.com 3.2.10. Victims The distribution of victims is confined only to Syria, but also reaches nearby countries. We have observed victims of the Syrian-based malware in:

- Syria
- Lebanon
- Turkey
- Kingdom of Saudi Arabia
- Egypt
- Jordan
- Palestine
- United Arab Emirates
- Israel
- Morocco
- United States

29 TLP: Green TLP: Green For any inquire please contact intelreports@kaspersky.com For any inquire please contact intelreports@kaspersky.com Below are snapshots taken from videos published by the attackers, showing their RAT control panel and list of victims. This shows some of the victims located in different countries. The sample details and domain lists used by the attackers can be found in Appendices 1 and 2 in the end of this document. 3.2.11. Activist Behavior It is worth noting that we have seen evidence of activists trying to carry out Denial of Service attacks on the RAT domains and servers, in an effort to overwhelm

their resources and cause their connections to timeout. The post below shows a warning from activists about pro-government hacker attacks on Facebook pages, explaining how pro-government groups post links to Trojanized applications in order to infect users. The activists announce in the post that they have spotted a C&C domain used by the Trojans and that they are attacking it to remove all hacked victims.²⁸ TLP: Green For any inquire please contact intelreports@kaspersky.com Victims geographical distribution map Map showing geographical distribution of victims with zoom on the most affected areas

GT: Kaspersky Lab detected the following types of RATs in MENA in 2013-2014: Zapchast, Bitfroze, Fynloski, and XtremeRAT.

A: There are several types of RATs that were detected in MENA in 2013-2014. These include NjRat, Backdoor.Win32.Bifrose, Backdoor.Win32.Fynloski, and Backdoor.Win32.XtremeRAT. Additionally, there are other types of RATs such as DarkComet and BlackShades that were also detected in MENA during this time period.

7. **Q:** What is a campaign? Which is the country with some of the most notorious actors in cyber attack campaigns? Why?

C: 6/1/2020 Russian Cyber Attack Campaigns and Actors - Threat Research

<https://ironnet.com/blog/russian-cyber-attack-campaigns-and-actors/> 1/9May 29, 2020 Russian Cyber Attack Campaigns and Actors ironnet.com /blog/russian-cyber-attack-campaigns-and-actors Some of the most notorious actors in the cyber threat landscape have been traced back to sponsorship by the Russian state. As the digital revolution has accelerated, so, too, has the Russian cyber attack landscape — hold-over Cold War tactics that evolved to take advantage of new electronic methods of communication. Strategic Russian interests are guided by the desires for Russia to be recognized as a great power, to protect the Russian identity, and to limit global United States power.^{xa0} These themes are evident in components commonly associated with Russian-backed cyber threat campaigns: The weaponization of information through disinformation campaigns and propaganda Attempted interference in democratic processes Strategic positioning within critical infrastructure, perhaps as preparation for potential escalation of hostilities with rival nations. To summarize the threat at a more tactical level, we have scoured cybersecurity reporting in order to prepare an overview of cyber threat actors

observed more recently, and to which evidence- based analysis has assigned the likelihood of Russian state-sponsorship as probable

GT: A campaign is a grouping of adversarial behaviors that describes a set of malicious activities or attacks that occur over a period of time against a specific set of targets. The country with some of the most notorious actors in cyber attack campaigns is Russia. Strategic Russian interests are guided by the desires for Russia to be recognized as a great power, to protect the Russian identity, and to limit global United States power. These themes are evident in components commonly associated with Russian-backed cyber threat campaigns: the weaponization of information through disinformation campaigns and propaganda; attempted interference in democratic processes; strategic positioning within critical infrastructure, perhaps as preparation for potential escalation of hostilities with rival nations.

A: A campaign is a series of coordinated actions or events intended to achieve a specific goal. In the context of cyber attacks, a campaign refers to a series of cyber attacks carried out by a group or individual with the aim of achieving a particular objective. The country with some of the most notorious actors in cyber attack campaigns is Russia. This is because Russia has been involved in several high-profile cyber attacks over the years, including the 2016 US presidential election hack, the NotPetya ransomware attack, and the SolarWinds supply chain attack. Russia is known for its sophisticated cyber capabilities and has been accused of using these capabilities to further its political and strategic objectives.

8. **Q:** What are SunBird-specific functionality?

C: Locally on the infected device, the data is collected in SQLite databases which are then compressed into ZIP files as they are uploaded to C2 infrastructure. SunBird can exfiltrate the following list of data, in addition to the list above:

List of installed applications

Browser history

Calendar information

BlackBerry Messenger (BBM) audio files, documents and images

WhatsApp Audio files, documents, databases, voice notes and images

Content sent and received via IMO instant messaging application

In addition to the list of actions above, SunBird can also perform the following actions:

Download attacker specified content from FTP shares

Run arbitrary commands as root, if possible

Scrape BBM messages and contacts via accessibility services

Scrape BBM notifications via accessibility services

_ Samples of SunBird have been found hosted on third-party app stores, indicating one possible distribution mechanism

GT: SunBird attempts to upload all data it has access to at regular intervals to its command and control (C2) servers. Locally on the infected device, the data is collected in SQLite databases which are then compressed into ZIP files as they are uploaded to C2 infrastructure.

SunBird can exfiltrate the following list of data:

-List of installed applications

-Browser history

-Calendar information

-BlackBerry Messenger (BBM) audio files, documents and images

-WhatsApp Audio files, documents, databases, voice notes and images

-Content sent and received via IMO instant messaging application

SunBird can also perform the following actions:

-Download attacker specified content from FTP shares

-Run arbitrary commands as root, if possible

-Scrape BBM messages and contacts via accessibility services

-Scrape BBM notifications via accessibility services

A: SunBird is a malicious Android spyware that was discovered by the Lookout Threat Intelligence team. It is designed to collect sensitive information from infected devices, including contact lists, call logs, text messages, and location data. The spyware also has the ability to record phone calls and take pictures or videos without the user's knowledge. One of the key features of SunBird is its ability to exfiltrate data from popular messaging apps such as WhatsApp, Telegram, and Signal. This allows the attackers behind the spyware to gain access to private conversations and other sensitive information shared through these platforms. Additionally, SunBird can also steal login credentials for various online accounts, allowing

the attackers to gain access to even more personal information. Overall, SunBird is a highly sophisticated piece of malware that poses a significant threat to users who may be targeted by its creators. It is important for users to be aware of this threat and take steps to protect themselves, such as installing antivirus software and avoiding downloading apps from untrusted sources

9. **Q:** What are the challenges that a defender deal with due to the complexity of vendors landscape?

C: Part II:

The defender landscape

Security Challenges and Opportunities for Defenders
The economic reality for service providers is that unless provided as a managed service, security is a cost center, not a profit center, and therefore needs to be kept lean—but the pressures of competition and the threat landscape have forced an increased focus on security. Service provider scale creates challenges As in every industry, the proliferation of security vendors and tools is a problem, because solutions are often not integrated and don't provide an actionable view of the threats facing the provider. In the service provider space, this problem is magnified by the scale of the market. Two-thirds of service provider security professionals said they rely on six or more vendors; 38 percent said they rely on more than 10 vendors (Figure 60). When asked about products in use, 70 percent said they use at least 6 security products, and half use more than 10 products. In many cases, say Cisco experts in this market, there isn't much integration among products, which means they experience an exponential increase in complexity for each incremental gain in security. Source: Cisco 2017 Security Capabilities Benchmark Study
Two-thirds of service providers use 6+ vendors, with 38% using more than 10 70% use 6+ products, with half using more than 10 38% 66% Service Providers
Products 70% 50% 6+ 10+ 6+ 10+ Vendors
Service Providers
Figure 60 Percentage of service providers that use solutions from 6 or more vendors and products
Download the 2017 graphics at: cisco.com/go/mcr2017graphics

GT: Defenders face several challenges due to the complexity of the vendor landscape:

- Managing multiple vendors and products adds significant complexity.
- The difficulty of integrating these products into a cohesive security strategy: they don't

provide an actionable view of the threats.

- Vulnerabilities due to the lack of coordination between different security tools.
- It is difficult for defenders to stay up-to-date on the latest security trends and best practices, leading to a potential loss of efficiency and effectiveness.

A: The challenges that a defender deals with due to the complexity of vendors landscape include the need to manage multiple vendors and products, the difficulty of integrating these products into a cohesive security strategy, and the potential for gaps in coverage or vulnerabilities due to the lack of coordination between different security tools. Additionally, the sheer number of vendors and products can make it difficult for defenders to stay up-to-date on the latest security trends and best practices, leading to a potential loss of efficiency and effectiveness

10. **Q:** Describe the Permanent Denial of Service attack (PDoS).

C: 40Cisco 2017 Midyear Cybersecurity Report

Attacker BehaviorBrickerBot

27 For more on this topic, see “BrickerBot PDoS Attack: Back With A Vengeance,” Radware, April 21, 2017: security.radware.com/ddos-threats-attacks/brickerbot-pdos-back-with-vengeance/ . 28 For more on this topic, see “Hajime – Sophisticated, Flexible, Thoughtfully Designed and Future-Proof,” by Pascal Geenens, Radware, April 26, 2017:

blog.radware.com/security/2017/04/hajime-futureproof-botnet/ .

Permanent denial of service (PDoS) attacks are fast-moving bot attacks designed to stop device hardware from functioning. This form of cyber attack is becoming increasingly popular.²⁷ Known as “phlashing” in some circles, PDoS attacks damage systems so severely that the hardware must be reinstalled or replaced. By exploiting security flaws or misconfigurations, PDoS attacks can destroy the firmware and basic system functions. BrickerBot can:

- Compromise devices: BrickerBot’s PDoS attacks use Telnet brute force—the same exploit vector used by Mirai— to breach users’ devices.
- Corrupt devices: Once it successfully accesses a device, BrickerBot performs a series of Linux commands that ultimately lead to corrupted storage. It then issues commands to disrupt Internet connectivity and device performance, wiping all files on the device.

Figure 33 shows the exact sequence of commands the BrickerBot performs.

Hajime

Hajime is intriguing, and threat intelligence researchers monitor it very closely. That's because it has not yet taken any action with the hundreds of thousands of devices it has so far infected. It is very large and, therefore, worrisome. The operator of Hajime claims to be a white hat hacker (Figure 34). Figure 33 Command sequence of BrickerBot

GT: Permanent denial of service (PDoS) attacks are rapid bot attacks that render device hardware non-functional, often requiring hardware reinstallation or replacement. These attacks, also known as "phlashing," exploit security flaws to destroy firmware and system functions. The BrickerBot malware uses Telnet brute force to breach devices, corrupts storage with Linux commands, disrupts internet connectivity, and wipes all files on the device.

A: The Permanent Denial of Service attack (PDoS) is a type of cyber attack that aims to render a computer or network unavailable for an extended period of time. Unlike other types of DDoS attacks, which aim to overwhelm a target with traffic, PDoS attacks seek to corrupt or destroy the target's hardware or software. This can be achieved through various means, such as exploiting vulnerabilities in the target's operating system or firmware, or by sending malicious code that causes the target to crash or become unresponsive. PDoS attacks can have serious consequences, including data loss, financial losses, and disruption of critical services. To prevent PDoS attacks, organizations should implement strong security measures, such as regular updates and patches, firewalls, and intrusion detection systems

11. **Q:** According to CWE what are the leading errors?

C: Vulnerability categories: Buffer errors remain in the lead In examining Common Weakness Enumeration (CWE) threat categories, buffer errors remain the most common type of coding error exploited by criminals (see Figure 41). This is a coding error repeatedly made by software developers. To prevent this error, developers should ensure that buffers are restricted so they can't be exploited Vulnerabilities

GT: Bufferoverflow is identified by Common Weakness Enumeration (CWE) as the most common type of coding error exploited by criminals highlighting the need for developers to restrict buffers to prevent exploitation.

A: The leading errors according to CWE are buffer errors. These are coding errors that are frequently made by software developers. To prevent these errors, developers should ensure that buffers are restricted so they cannot be exploited.

12. **Q:** What are the systemic issues or mistakes made by defenders that contributed to attackers' success?

C: 48Recommendations For Defenders

While reviewing 2023's incidents, we kept in mind two topics close to defenders' hearts; factors that contributed to attackers' success, and mitigations defenders can use to frustrate attackers. This section distills the most critical lessons into actionable recommendations for defenders. These are pragmatic, effective approaches to counter emerging and evolving cyberthreats18Incident Response Report 2024 18 Spotlight On: Speed

Clearly, defenders need to speed up as much as possible. We'll give more recommendations for how to do that later in this report, but for the moment, there are a few key principles to keep in mind:

- **Preparation.** One of the best ways to get ahead of attackers is to truly get ahead. Through proactive preparation, organizations can get ready to respond within hours to compromises, stopping attackers before they have a chance to execute their plans.
- **Automation.** As we mentioned in the Black Basta example above, human security teams need to sleep, and they're also typically pulled in many directions. Automation—especially when powered by machine learning and AI—can help defenders sift through alerts and surface the ones that truly need attention.
- **Zero Trust.** Another powerful way to limit the damage an attacker can do after compromise is to limit their movement and activity. When organizations design their security posture in accordance with a Zero Trust philosophy, attackers become less powerful when they gain initial access, because initial access means much less. It's the difference between a thief getting into an entryway and being able to move through the entire building, and a thief getting into the lobby only to encounter another locked door.
- **Defense in depth.** A security program designed with overlapping defenses and controls gives attackers more opportunity to alert you to their presence. Especially combined with limited

privileges in a Zero Trust philosophy, you can raise the signal-to-noise ratio of meaningful alerts that will let you focus on attacker activity earlier in the attack lifecycle.

GT: Systemic issues or mistakes by defenders that contribute to attackers' success include:

- Lack of Preparation: Failing to proactively prepare for compromises results in delayed responses.
- Insufficient Automation: Over-reliance on manual processes can lead to missed critical alerts and slower response times.
- Weak Zero Trust Implementation: Not restricting attackers' movement after initial access allows them to cause more damage.
- Inadequate Defense in Depth: Single-layer defenses without overlapping controls make it easier for attackers to navigate and compromise systems.

A: The systemic issues or mistakes made by defenders that contributed to attackers' success include: 1. Lack of preparedness: Defenders often fail to prepare adequately for potential attacks, leaving them vulnerable to sophisticated attacks. 2. Slow response time: Defenders take too long to detect and respond to attacks, giving attackers ample time to carry out their malicious activities. 3. Insufficient resources: Defenders may not have enough resources to effectively monitor and protect their systems, making it easier for attackers to exploit vulnerabilities. 4. Poor communication: Defenders may not communicate effectively with each other or with external parties, leading to delays in incident response and increased risk of data breaches. 5. Inadequate training: Defenders may lack the necessary skills and knowledge to identify and respond to complex attacks, leaving them vulnerable to attackers who are better equipped. 6. Overconfidence: Defenders may underestimate the capabilities of attackers, leading to complacency and a false sense of security. To mitigate these issues, defenders should prioritize preparation, invest in automation and AI-powered tools, implement a Zero Trust architecture, adopt defense-in-depth strategies, improve communication and collaboration, and provide ongoing training and education to their teams.

13. **Q:** What are Tactics, Techniques and Procedures (TTPs)? How they are linked with the concept of 'exploit target'?

C: tk — Appendix B — Mitre ATT&CK; Techniques

*Tactic: Technique** | **ATT&CK; Code** —|—

*Resource Development:** Acquire Infrastructure: Virtual Private Server | T1583.003

*Resource Development:** Compromise Infrastructure: Server | T1584

*Reconnaissance:** Active Scanning: Vulnerability Scanning | T1595CYBER

THREAT ANALYSIS Appendix B — Mitre ATT&CK Techniques Tactic: Technique ATT&CK

Code Resource Development: Acquire Infrastructure: Virtual Private Server T1583.003 Re-

source Development: Compromise Infrastructure: Server T1584 Reconnaissance:Active Scan-

ning: Vulnerability Scanning T1595

GT: Tactics, Techniques and Procedures (TTP) are representations of the behavior or modus operandi of cyber adversaries. They can include malware, social engineering techniques, and other means of gaining access to systems or data. Exploit targets are vulnerabilities or weaknesses in software, systems, networks or configurations that are targeted for exploitation by the TTP.

A: Tactics, Techniques and Procedures (TTPs) refer to the methods and techniques used by threat actors to carry out cyber attacks. They include the tools, techniques, and procedures employed by attackers to gain access to systems, exfiltrate data, or cause damage. The concept of 'exploit target' refers to the specific vulnerabilities or weaknesses in a system that can be exploited by attackers using TTPs. By understanding the TTPs and exploit targets, organizations can better defend against cyber threats and mitigate potential risks.

14. **Q:** Provide a definition of an 'attack'. What type of supply chain attack occurred in September 2017?

C: 4 Ibid. 5 Ukraine scrambles to contain new cyber threat after 'NotPetya' attack, by Jack Stubbs and Matthias Williams, Reuters, July 2017: [reuters.com/article/us-cyber-attack-ukraine-backdoor/ukraine-scrambles-to-contain-new-cyber-threat-after-notpetya-attack-idUSKBN19Q14P](https://www.reuters.com/article/us-cyber-attack-ukraine-backdoor/ukraine-scrambles-to-contain-new-cyber-threat-after-notpetya-attack-idUSKBN19Q14P)

. 6 The MeDoc Connection, Cisco Talos blog, July 2017: blog.talosintelligence.com/2017/07/the-medoc-connection.htm 1.

7 CCleaner Command and Control Causes Concern, Cisco Talos blog, September 2017: blog.talosintelligence.com/2017/09/ccleaner-c2-concern.html .Security weak spot: the supply chain The Nyetia campaign was also a supply chain attack, one of

many that Cisco threat researchers observed in 2017. Targeted attacks and insider exfiltration are biggest concerns. In keeping with concerns about revenue loss and brand damage, retail security professionals said targeted attacks (38 percent) and insider exfiltration (32 percent) pose the highest security risks to their organizations (Figure 64). They are right to be concerned: Often, attacks begin inside an organization. That means security built around examining indicators of compromise (IOCs) isn't enough. Organizations also need tools for reviewing indicators of attacks. To detect sophisticated targeted attacks, like APTs or phishing attacks, retailers need to distinguish between normal and abnormal traffic patterns, which can vary by day, week, or shopping season. The Federal Energy Regulatory Commission (FERC) recently directed the North American Energy Reliability Corporation (NERC) to develop new standards for critical infrastructure protection, specifically directed at the utility supply chain. The standards are expected to address supply chain risk management for industrial control system hardware, software, and computing and networking services associated with bulk electric system operations.⁵² Targeted attacks and APTs are key concern. Targeted attacks are high on the list of worries for utility and energy security professionals. Security professionals said targeted attacks (42 percent) and advanced persistent threats, or APTs (40 percent) were the most critical security risks to their organizations (Figure 68).

GT: An attack is the use of an exploit(s) by an adversary to take advantage of a weakness(s) with the intent of achieving a negative technical impact(s). In September 2017, a supply chain attack occurred in which a software vendor distributed a legitimate software package known as CCleaner, but the binaries contained a Trojan backdoor that was signed with a valid certificate, giving users false confidence that the software they were using was secure. The actors behind this campaign were targeting major technology companies where the software was in use, either legitimately or as part of shadow IT.

A: A supply chain attack is a type of cyber attack that targets the supply chain of a company or organization. In this case, the attack occurred in September 2017 when a malicious code was inserted into a popular software update tool called CCleaner. This tool is used by millions of people worldwide to clean up their computers and remove unwanted files. However, the version of CCleaner that was released in September 2017 contained a backdoor that allowed hackers to gain access to the systems of users who installed the update. This attack was

particularly significant because it affected a large number of users and could have been used to steal sensitive information or disrupt critical systems

E Cybersecurity RAGOnto queries

1. **Question:** What is signature detection used for in antivirus software?

These concepts might be useful to answer the question:

RogueSecuritySoftware, Malware, Vulnerability, Software, Application, Browser, OperatingSystem, VirusScanner, System, GeneralVirus, VirusScanner

Context: Antivirus software uses either signature or heuristic-based detection to identify malware. On the one hand, signature detection searches for a unique sequence of bytes that is specific to a piece of malicious code

Ground Truth: the same as the cybersecurity RAG one

Answer: Antivirus software uses signature detection to identify malware. Signature detection looks for a unique sequence of bytes that is specific to a piece of malicious code. This allows the software to detect and remove the malware from the system.

2. **Q:** Which sectors and organizations did RedJuliatt target in their cyber attacks in and outside Taiwan?

These concepts might be useful to answer the question:

Organization, Country, Location

C: THREAT ANALYSIS

RedJuliatt compromised government organizations in Taiwan, Laos, Kenya, and Rwanda. The group also targeted the technology industry in Taiwan, including an optoelectronics company, a large Taiwanese facial recognition company that has held contracts with the Taiwanese government, and four software companies. Multiple universities were also compromised, including three Taiwanese universities, an American university, and a Djiboutian university. The group also targeted religious organizations in Taiwan, Hong Kong, and South Korea. Additionally, RedJuliatt compromised a geological engineering company in Hong Kong, a Taiwanese waste

and pollution treatment company, and a Taiwanese publishing house.

Figure 3x00

Suspected RedJuliett victims by sector x00Source: Recorded Future)

RedJuliett Targets Internet-Facing Web Servers and Appliances Using Recorded Future® Network Intelligence, Insikt Group observed RedJuliett using dedicated threat actor-controlled SoftEther VPN servers to conduct reconnaissance and attempted exploitation activity targeting a wide range of predominantly Taiwanese organizations from November 2023 to April 2024. In addition to acting as SoftEther VPN nodes, multiple RedJuliett IP addresses have been concurrently configured as Acunetix Web Application Security Scanners. Insikt Group has regularly observed vulnerability scanning and exploitation activity from these servers targeting web servers associated

6 C T A x00 C N x00 2 0 2 4 x00 0 6 2 4 R e c o r d e d F u t u r e ® | w w w . r e c o r d e d f u t u r e . c o m

GT: the same as the cybersecurity RAG one

A: Organizations in Hong Kong, Malaysia, Laos, South Korea, the United States, Djibouti, Kenya and Rwanda.

3. **Q:** How do pro-government groups use Trojanized apps on Facebook to infect users, and what techniques do they use for delivery?

These concepts might be useful to answer the question: RemoteAccessTrojan-RAT, Malware, Vulnerability, Trojan, Application, Software, System, RemoteAccessTrojan-RAT, Phishing, Infrastructure

C: 29TLP: Green TLP: Green For any inquire please contact intelreports@kaspersky.com For any inquire please contact intelreports@kaspersky.com

Below are snapshots taken from videos published by the attackers, showing their RAT control panel and list of victims. This shows some of the victims located in different countries. The sample details and domain lists used by the attackers can be found in Appendices 1 and 2 in the end of this document.

3.2.11. Activist Behavior

It is worth noting that we have seen evidence of activists trying to carry out Denial of Service

attacks on the RAT domains and servers, in an effort to overwhelm their resources and cause their connections to timeout. The post below shows a warning from activists about pro-government hacker attacks on Facebook pages, explaining how pro-government groups post links to Trojanized applications in order to infect users. The activists announce in the post that they have spotted a CC domain used by the Trojans and that they are attacking it to remove all hacked victims.⁶ TLP: Green For any inquire please contact intelreports@kaspersky.com

3. Analysis

3.1. Infection Vectors

Malware writers are using multiple techniques to deliver their files and entice the victims to run them, creating an effective infection vector. Mainly depending on social engineering the attackers exploit:

- Victims' trust in social networking forums
- Victims' curiosity in following news related to political conflict in Syria
- Victims' fear of attacks from government
- Victims' lack of technology awareness

Once they have infected the victim's computer, attackers have full access and control over victim's devices. In the following section we show different versions of posts sent via popular file sharing sites or social networking platforms. The sample details and domain lists used by the attackers can be found in the Appendices 1 and 2 in the end of the document.

3.1.1. Skype messages

Messages sent via Skype offer links to download: 1.The "SSH VPN" program to encrypt communication

2.The popular and effective antivirus with daily updates from "Ammazon Internet Security"

3.The "SmartFirewall" to block connections made by malware and bad programs. The messages are usually sent from fake or compromised accounts.

5TLP: Green TLP: Green For any inquire please contact intelreports@kaspersky.com For any inquire please contact intelreports@kaspersky.com

2. Introduction

The geopolitical conflicts in the Middle East have deepened in the last few years; Syria is no exception. The crisis is taking many forms, and the cyberspace conflict is intensifying as

sides try to tilt the struggle, by exploiting cyber intelligence and exercising distortion. In the last few years cyber-attacks in Syria have moved into the front line; many activities in cyberspace have been linked to Syria, especially those conducted by the Syrian Electronic Army and pro-government groups. The Global Research and Analysis Team (GReAT) at Kaspersky Lab has found new malware attacks in Syria, using new but not advanced techniques to hide and operate malware, in addition to using proficient social engineering tricks to deliver malware by tricking and tempting victims into opening and exploring malicious files. The malware files have been found on hacked activist sites, web pages and in social networking forums. Cyber Arabs, an Arabic-language digital security project of the IWPR (Institute for War and Peace Reporting), reported four of these samples in March 2014. The same samples were also reported on Syrian Facebook pages (Technicians For Freedom): <https://www.facebook.com/tech4freedom> Given the complexity of the situation, there are many factors and entities at play in this event, but from the outside these are all largely speculative. Pro-government groups talk about “defense” and opposition activists talk about “offense”. Here, we will only focus on the malware and the facts that have been found during the analysis, presenting only relevant information, in the hope of setting a clear context for this research.

TLP: Green For any inquire please contact intelreports@kaspersky.com

Contents	1.	Executive Summary	2.	Introduction	5
	3.	Analysis	6	3.1. Infection Vectors	6
	3.1.1.	Skype messages	6	3.1.2. Facebook posts	7
	3.1.3.	YouTube Videos	8	3.2. Samples and types of files	9
	3.2.1.	The National Security Program	9	3.2.2. Files named “Scandals” are quite attractive	14
	3.2.3.	“Ammazon Internet Security” the “popular Antivirus”	16	3.2.4. You’ve installed the latest antivirus solution, now let’s “protect your network”	19
	3.2.5.	Whatsapp and Viber for PC: Instant messaging, instant infection	20	3.2.6. Beware of chemical attacks	22
	3.2.7.	Commands and functionality	23	3.2.8. Evolution of malware attack file numbers	25
	3.2.9.	Locations, domains and team	26	3.2.10. Victims	28
	3.2.11.	Activist Behavior	30	3.3. Attribution	32
	4.	Kaspersky Lab MENA RAT Statistics	34	5.	Conclusion
		Appendix 1: Samples	38		
		Appendix 2: C&C Domains	47		

TLP: Green For any inquire please contact intelreports@kaspersky.com

TLP: Green For any inquire please contact intelreports@kaspersky.com The presence of DarkComet’s “DC_MUTEX-*” was a giveaway of the usage of this remote administration

tool. During infection, the Excel spreadsheet is displayed, comprising 96763 rows and 13 columns of activist information. The rows correspond to records of individuals wanted by the government and the columns correspond to information about the individuals. While there is no column description, data in each column reflects the type of data.

3.2.2. Files named “Scandals” are quite attractive

Using shockingly disturbing videos to distribute malware A disturbing video showing injured victims of recent bombings was used to appeal to people’s fear and exert them to download a malicious application available in a public file-sharing website. After our initial analysis, the file named “exe” proved to be heavily obfuscated with the commercial utility “MaxToCode” for .NET as a means of avoiding early detection by antivirus solutions. When executed, the original sample created another executable file in the Windows’ temporary folder (C:

Users
USERNAME

AppData

Local

Temp) named “Trojan.exe”, which corresponds to the code of the RAT itself. This is used to save all keystrokes and system activity to another file in the same location, “Trojan.exe.tmp”.⁷Incident Response Report 2024 Who’s Doing It? In 2023, by far the most damaging threat actor was a group we call Muddled Libra. They’re a criminal group focused on /uniFB01nancial gain. They are also one of the most aggressive threat actors in the landscape today. Countering their attacks is a challenge for defenders of small and large organizations alike. This report has a special section on them, including an analysis of what they do and how you can defend against them. Of course, they weren’t the only group in the mix. We also responded to compromises from other criminal groups, state-aligned actors, and some groups that we don’t know enough about yet to characterize. Answering this question of “who” is one of the things that Unit 42 Threat Intelligence is great at. We’ll keep researching and uncovering attackers wherever they try to hide. How Do Threat Actors Get In? In the past year, attackers’ initial access most often started with a software vulnerability. The largest attack campaigns began with successful exploitation of internet-facing systems. From there, the stories varied. Naturally, we also saw other techniques, from attackers using previously stolen credentials,

phishing (both in email and chat), and more. But don't be misled—though phishing lost the top spot and dropped to number three, it's still a serious threat. Attackers leverage phishing attacks for access, credentials, and tokens, rather than trying to drop malware. What Helped Threat Actors Succeed? There were three security weaknesses that weren't themselves the cause of incidents, but made attackers' lives easier. Address these, and you'll be helping yourself and hurting them.

- Patch management. It's only a matter of time. Unpatched vulnerabilities on internet-facing systems will be exploited. Measure and reduce your attack surface, and you'll improve this.
- Consistent coverage. Our incident responders found that organizations with partial or incomplete deployment of security controls (especially endpoint detection and response tools) allowed attackers to operate from parts of the network that weren't defended. Deploy your defenses everywhere, and you'll deny them this advantage.
- Identity and access management. More attackers are stealing and using the identity of authorized personnel to access and move around networks.

Make that harder and embark on or continue a journey toward Zero Trust networking, and you'll pressure attackers into making mistakes you can see. Actors, Methods and Targets Subscribe Now Listen to Unit 42's Where we share unique threat intelligence insights, new threat actor TTPs, and real-world case studies

GT: the same as the cybersecurity RAG one

A: Remote Access Trojans (RATs) are a type of malware that allows attackers to remotely access and control a victim's computer. These trojans are typically delivered through phishing emails or drive-by downloads, where the user is tricked into downloading and running the malicious software. Once installed, the RAT provides the attacker with complete control over the victim's machine, allowing them to steal sensitive information, install additional malware, and even take control of the camera and microphone. One way that pro-government groups use Trojanized apps on Facebook to infect users is by creating fake profiles and posting links to malicious apps on their walls. When users click on the link, they are redirected to a malicious website where they are prompted to download and install the app. The app appears to be legitimate, but in reality, it contains a hidden RAT that gives the attacker full control over the victim's device. Another technique used by pro-government groups is to create fake news

articles or videos that contain malicious links. These links may appear to lead to legitimate news sources or videos, but instead, they redirect the user to a malicious website where they are prompted to download and install the malicious app. In addition to using Trojanized apps, pro-government groups also use phishing emails to infect users. These emails may appear to come from trusted sources such as banks or government agencies, and they may contain links to malicious websites or attachments that contain malware. Overall, pro-government groups use a variety of techniques to infect users with RATs, including Trojanized apps, phishing emails, and fake news articles or videos. By understanding these techniques, users can take steps to protect themselves, such as being cautious when clicking on links or downloading apps, and keeping their antivirus software up to date.

4. **Q:** What are RemoteAccessTrojan-RAT? What types of RATs did Kaspersky Lab detect in MENA in 2013-2014?

These concepts might be useful to answer the question:

RemoteAccessTrojan-RAT, Malware, Vulnerability

C: 33 TLP: Green TLP: Green For any inquire please contact intelreports@kaspersky.com

For any inquire please contact intelreports@kaspersky.com

4. Kaspersky Lab MENA RAT Statistics

Remote Administration Tool (RAT) Trojans are malicious programs that allow a remote “operator” to control a system as if he has physical access to that system. Malicious RATs are widely used by different types of cybercriminals (hacktivists, script-kiddies, and scammers) and even in some state-sponsored attacks. Some of the most popular RATs are detected by Kaspersky products as following:

- Trojan.MSIL.Zapchast, also known as Njrat
- Backdoor.Win32.Bifrose, also known as Bitfrose
- Backdoor.Win32.Fynloski, also known as DarkComet
- Backdoor.Win32.Xtreme, also known as Xtremrat

The statistics below, extracted from the Kaspersky Security Network (KSN), show the number of RAT infection attacks blocked by Kaspersky Lab products in the MENA (Middle East North Africa) region in the 2013-2014 period: Country/Detection Zapchast Bitfrose Fyn-

loski XtremeRAT Total Algeria 39113 12071 11643 7106 69900+ Turkey 6326 3325 14002 3586 27200+ KSA 9616 5555 5336 4516 25000+ Egypt 5567 5883 4325 2634 18400+ Iraq 6756 2280 3235 3055 15300+ UAE 3594 1165 9244 745 14700+ Morocco 4084 2710 3104 1233 11100+ Lebanon 426 297 8073 136 8900+ Tunisia 2844 1888 1495 1004 7200+ Syria 2806 1897 1362 544 6600+ Qatar 1332 327 2177 233 4000+ Jordan 1259 680 1104 414 3400+ Oman 1241 446 915 374 2900+ Bahrain 1218 178 1214 254 2800+ Kuwait 454 407 922 345 2100+

GT: the same as the cybersecurity RAG one

A: Remote Access Trojan (RAT) is a type of malware that allows attackers to remotely control a victim's computer or mobile device. It is a form of spyware, designed to allow a remote user to monitor and control the infected system as if they were physically present. It can be used to steal sensitive information, spy on users, and carry out other malicious activities. In 2013-2014, Kaspersky Lab detected several types of RATs in the Middle East and North Africa (MENA) region. These included Trojan.MSIL.Zapchast, Backdoor.Win32.Bifrose, Backdoor.Win32.Fynloski, and Backdoor.Win32.DarkComet. These RATs were used by various threat actors to target individuals and organizations in the region.

5. **Q:** What are threat actors? How do they exploit Skype messages?

These concepts might be useful to answer the question:

ThreatActor, Malware, Vulnerability, File, Indicator, Filepath

C: Skype messages

Messages sent via Skype offer links to download: 1.The "SSH VPN" program to encrypt communication

2.The popular and effective antivirus with daily updates from "Ammazon Internet Security"

3.The "SmartFirewall" to block connections made by malware and bad programs

The messages are usually sent from fake or compromised accounts.

Threat Actor: Threat Actors are actual individuals, groups, or organizations believed to be operating with malicious intent. A Threat Actor is not an Intrusion Set but may support or be affiliated with various Intrusion Sets, groups, or organizations over time.

Type - Person, Organization, Group Added - 06/2021

It is a weakness or defect in the requirements, designs, or implementations of the computa-

tional logic (e.g., code) found in software and some hardware components (e.g., firmware) that can be directly exploited to negatively impact the confidentiality, integrity, or availability of that system. (ref: STIX 2.1)

- Adapted from STIX 2.1 - Added 06/2021 - MalONT v2.0

GT: the same as the cybersecurity RAG one

A: Threat actors are individuals, groups, or organizations who engage in cyberattacks with malicious intent. They use various methods to exploit vulnerabilities in systems, including social engineering, phishing attacks, and malware distribution. One way threat actors exploit Skype messages is by sending malicious links or attachments that appear legitimate but contain malware. These links or attachments can lead to the installation of malware on the victim's device, which can then be used to steal sensitive information or take control of the device.

6. **Q:** What types of RATs did Kaspersky Lab detect in MENA in 2013-2014?

These concepts might be useful to answer the question:

RemoteAccessTrojan-RAT, Malware, Vulnerability

C: Some of the most popular RATs are detected by Kaspersky products as following:

- Trojan.MSIL.Zapchast, also known as Njrat
- Backdoor.Win32.Bifrose, also known as Bitfrose
- Backdoor.Win32Kaspersky Lab MENA RAT Statistics

Remote Administration Tool (RAT) Trojans are malicious programs that allow a remote “operator” to control a system as if he has physical access to that systemcom Based on KSN world statistics, the MENA region has one of the highest numbers for RAT attacks, as shown below:

Country Number of users

Algeria 39113

India 35024

France 10955

Saudi Arabia 9616

Mexico 6862

Iraq 6756

Turkey 6321

Egypt 5567

Russian Federation 5526

Malaysia 5014

- Algeria has the highest number of users facing NjRat infection for the 2013-2014 period and five countries from MENA are in the NjRat top 10

- Algeria has the highest number of users facing Xtreme RAT infection for the 2013-2014 period and four countries from MENA are in the Xtreme RAT top 10

Among the most popular RAT found in the samples subset is Dark Comet, a free remote administration tool that provides a comprehensive command set for the attackers to use in their malicious purposes. Xtreme, also known as Xtremrat. The statistics below, extracted from the Kaspersky Security Network (KSN), show the number of RAT infection attacks blocked by Kaspersky Lab products in the MENA (Middle East North Africa) region in the 2013-2014 period:

Country	Detection	Zapchast	Bitfrose	Fynloski	XtremeRAT	Total
Algeria	39113	12071	11643	7106	69900+	
Turkey	6326	3325	14002	3586	27200+	
KSA	9616	5555	5336	4516	25000+	
Egypt	5567	5883	4325	2634	18400+	
Iraq	6756	2280	3235	3055	15300+	
UAE	3594	1165	9244	745	14700+	
Morocco	4084	2710	3104	1233	11100+	
Lebanon	426	297	8073	136	8900+	
Tunisia	2844	1888	1495	1004	7200+	
Syria	2806	1897	1362	544	6600+	
Qatar	1332	327	2177	233	4000+	
Jordan	1259	680	1104	414	3400+	
Oman	1241	446	915	374	2900+	
Bahrain	1218	178	1214	254	2800+	
Kuwait	454	407	922	345	2100+	

Malicious RATs are widely used by different types of cybercriminals (hacktivists, script-kiddies, and scammers) and even in some state-sponsored attacks. Another RAT widely used in the Arab world is NjRAT, which includes a list of commands (see below) that can be sent from the controller to the infected system.

DarkComet Control panel & Functionality

Below are snapshots taken from videos published by the attackers, showing their RAT control panel and list of victims. This shows some of the victims located in different countries.

Known Targets

- Ukrainian financial sector
- Sample TTPs
- Spearphishing with Microsoft Excel attachments containing malicious macros
- Hiding malicious network activity by abusing legitimate services to host payloads or provide communication mediums for attackers
- Deploying redundant back-

doors within a network Disguising malware backdoors by providing them with names resembling A V-related services Also Known AsSandworm T eam Energetic Bear This group is assessed as the creator of the Havex RAT, which is one of five known ICS tailored malware familiesWhile silently executing a remote administration tool when launching this “security suite”, targeted victims were left without their “Ammazon” protection but with a RAT installed

Samples Analysis has led us to identify the following RAT variants being used in the wild:

- ShadowTech RAT
- Xtreme RAT
- NjRAT
- Bitcomet RAT
- Dark Comet RAT
- BlackShades RAT

The samples collected during our research can be classified as follows

GT: the same as the cybersecurity RAG one

A: The most popular RATs detected by Kaspersky Lab in MENA in 2013-2014 include Zapchast, Bitfrose, Fynloski, Xtreme RAT, and NjRat. These RATs have been used by various cybercriminals, including hackers, script-kiddies, and scammers. Some of these RATs, such as NjRAT, have a list of commands that can be sent from the controller to the infected system.

7. **Q:** What is a campaign? Which is the country with some of the most notorious actors in cyber attack campaigns? Why?

These concepts might be useful to answer the question:

Campaign, Vulnerability, Country, Location, AttackPattern, AttackPattern_LargeDescription, AttackPattern_SmallDescription, AttackPattern_LargeDescription, AttackPattern_SmallDescription, ThreatActor, Malware

C: 6/1/2020 Russian Cyber Attack Campaigns and Actors - Threat Research

<https://ironnet.com/blog/russian-cyber-attack-campaigns-and-actors/> 1/9May 29, 2020

Russian Cyber Attack Campaigns and Actors

[ironnet.com /blog/russian-cyber-attack-campaigns-and-actors](https://ironnet.com/blog/russian-cyber-attack-campaigns-and-actors)

Some of the most notorious actors in the cyber threat landscape have been traced back to sponsorship by the Russian state. As the digital revolution has accelerated, so, too, has the Russian cyber attack landscape — hold-over Cold War tactics that evolved to take advantage of new electronic methods of communication. Strategic Russian interests are guided by the desires for Russia to be recognized as a great power, to protect the Russian identity, and to limit global United States power. These themes are evident in components commonly associated with Russian-backed cyber threat campaigns: The weaponization of information through disinformation campaigns and propaganda Attempted interference in democratic processes Strategic positioning within critical infrastructure, perhaps as preparation for potential escalation of hostilities with rival nations. To summarize the threat at a more tactical level, we have scoured cybersecurity reporting in order to prepare an overview of cyber threat actors observed more recently, and to which evidence- based analysis has assigned the likelihood of Russian state-sponsorship as probable

GT: the same as the cybersecurity RAG one

A: A campaign is a series of coordinated actions or events intended to achieve a specific goal. In the context of cyber attacks, a campaign refers to a series of cyber attacks carried out by a group or individual with the aim of achieving a particular objective. The country with some of the most notorious actors in cyber attack campaigns is Russia. This is because Russia has been involved in several high-profile cyber attacks over the years, including the 2016 US presidential election hack, the NotPetya ransomware attack, and the SolarWinds supply chain attack. Russia is known for its sophisticated cyber capabilities and has been accused of using these capabilities to further its political and strategic objectives.

8. **Q:** What are SunBird-specific functionality?

These concepts might be useful to answer the question:

HostingMalware, Infrastructure, Malware, Vulnerability, Adware, Backdoor, Bootkit, Bot, DDoS, Downloader, Dropper, Exploit-kit, GeneralVirus, Keylogger, Ransomware, RemoteAccessTrojan-RAT, RogueSecuritySoftware, Rootkit, Spyware, Trojan, Webshell, ResourceExploitation, ScreenCapture, Unknown, Wiper, Worm, MalwareAnalysis, MalwareFamily

C: Locally on the infected device, the data is collected in SQLite databases which are then

compressed into ZIP files as they are uploaded to C2 infrastructure. SunBird can exfiltrate the following list of data, in addition to the list above:

List of installed applications

Browser history

Calendar information

BlackBerry Messenger (BBM) audio files, documents and images

WhatsApp Audio files, documents, databases, voice notes and images

Content sent and received via IMO instant messaging application

In addition to the list of actions above, SunBird can also perform the following actions:

Download attacker specified content from FTP shares

Run arbitrary commands as root, if possible

Scrape BBM messages and contacts via accessibility services

Scrape BBM notifications via accessibility services

_Samples of SunBird have been found hosted on third-party app stores, indicating one possible distribution mechanism

GT: the same as the cybersecurity RAG one

A: SunBird is a malicious Android spyware that was discovered by the Lookout Threat Intelligence team. It is designed to collect sensitive information from infected devices, including contact lists, call logs, text messages, and location data. The spyware also has the ability to record phone calls and take pictures or videos without the user's knowledge. One of the key features of SunBird is its ability to exfiltrate data from popular messaging apps such as WhatsApp, Telegram, and Signal. This allows the attackers behind the spyware to gain access to private conversations and other sensitive information shared through these platforms. Additionally, SunBird can also steal login credentials for various online accounts, allowing the attackers to gain access to even more personal information. Overall, SunBird is a highly sophisticated piece of malware that poses a significant threat to users who may be targeted by its creators. It is important for users to be aware of this threat and take steps to protect themselves, such as installing antivirus software and avoiding downloading apps from untrusted sources.

9. **Q:** What are the challenges that a defender deal with due to the complexity of vendors landscape?

These concepts might be useful to answer the question:

C: Part II:

The defender landscape

Security Challenges and Opportunities for Defenders
The economic reality for service providers is that unless provided as a managed service, security is a cost center, not a profit center, and therefore needs to be kept lean—but the pressures of competition and the threat landscape have forced an increased focus on security. Service provider scale creates challenges. As in every industry, the proliferation of security vendors and tools is a problem, because solutions are often not integrated and don't provide an actionable view of the threats facing the provider. In the service provider space, this problem is magnified by the scale of the market. Two-thirds of service provider security professionals said they rely on six or more vendors; 38 percent said they rely on more than 10 vendors (Figure 60). When asked about products in use, 70 percent said they use at least 6 security products, and half use more than 10 products. In many cases, say Cisco experts in this market, there isn't much integration among products, which means they experience an exponential increase in complexity for each incremental gain in security. Source: Cisco 2017 Security Capabilities Benchmark Study
Two-thirds of service providers use 6+ vendors, with 38% using more than 10
70% use 6+ products, with half using more than 10
38% 66% Service Providers
Products 70% 50% 6+ 10+ 6+ 10+ Vendors
Service Providers
Figure 60 Percentage of service providers that use solutions from 6 or more vendors and products
Download the 2017 graphics at: cisco.com/go/mcr2017graphics

GT: the same as the cybersecurity RAG one

A: The challenges that a defender deals with due to the complexity of vendors landscape include the need to manage multiple vendors and products, the difficulty of integrating these products into a cohesive security strategy, and the potential for gaps in coverage or vulnerabilities due to the lack of coordination between different security tools. Additionally, the sheer number of vendors and products can make it difficult for defenders to stay up-to-date on the latest security trends and best practices, leading to a potential loss of efficiency and

effectiveness.

10. **Q:** Describe the Permanent Denial of Service attack (PDoS).

These concepts might be useful to answer the question:

DenialOfService, ExcessForks, MemoryConsumption, NetFlood, PingOfDeath, SysCrash, Consequence, MeansOrConsequence

C: 40 Cisco 2017 Midyear Cybersecurity Report

Attacker BehaviorBrickerBot

27 For more on this topic, see “BrickerBot PDoS Attack: Back With A Vengeance,” Radware, April 21, 2017:

security.radware.com/ddos-threats-attacks/brickerbot-pdos-back-with-vengeance/ .

28 For more on this topic, see “Hajime – Sophisticated, Flexible, Thoughtfully Designed and Future-Proof,” by Pascal Geenens, Radware, April 26, 2017: blog.radware.com/security/2017/04/hajime-futureproof-botnet/ .

Permanent denial of service (PDoS) attacks are fast-moving bot attacks designed to stop device hardware from functioning. This form of cyber attack is becoming increasingly popular.²⁷ Known as “phlashing” in some circles, PDoS attacks damage systems so severely that the hardware must be reinstalled or replaced. By exploiting security flaws or misconfigurations, PDoS attacks can destroy the firmware and basic system functions. BrickerBot can:

- **Compromise devices:** BrickerBot’s PDoS attacks use Telnet brute force—the same exploit vector used by Mirai— to breach users’ devices.
- **Corrupt devices:** Once it successfully accesses a device, BrickerBot performs a series of Linux commands that ultimately lead to corrupted storage. It then issues commands to disrupt Internet connectivity and device performance, wiping all files on the device.

Figure 33 shows the exact sequence of commands the BrickerBot performs.

Hajime

Hajime is intriguing, and threat intelligence researchers monitor it very closely

GT: the same as the cybersecurity RAG one

A: The Permanent Denial of Service attack (PDoS) is a type of cyber attack that aims to render a computer or network unavailable for an extended period of time. Unlike traditional

denial of service attacks, which are typically short-lived and intended to cause temporary disruption, PDoS attacks are designed to permanently disable the target system. One common method of carrying out a PDoS attack is through the use of malicious software known as "brickers." These programs are specifically designed to corrupt or delete critical system files, rendering the target system unusable. In some cases, brickers may also attempt to overwrite the boot sector of the hard drive, making it impossible for the system to start up at all. Another technique used in PDoS attacks is the exploitation of vulnerabilities in the target system's firmware. This can allow attackers to gain remote access to the system and make changes to its configuration, potentially causing it to become unresponsive or even crash entirely. Once a PDoS attack has been successful, the affected system may require extensive repairs or replacement in order to restore functionality. As such, these attacks can have significant financial and operational impacts on organizations that rely on their IT infrastructure for day-to-day operations. To protect against PDoS attacks, organizations should implement robust security measures such as firewalls, intrusion detection systems, and regular patch management practices. Additionally, it is important to maintain strong password policies and ensure that all user accounts have appropriate levels of access control.

11. **Q:** According to CWE what are the leading errors?

These concepts might be useful to answer the question:

Vulnerability, BufferOverflow, DenialOfService, Means, Consequence, MeansOrConsequence

C: Vulnerability categories: Buffer errors remain in the lead In examining Common Weakness Enumeration (CWE) threat categories, buffer errors remain the most common type of coding error exploited by criminals (see Figure 41). This is a coding error repeatedly made by software developers. To prevent this error, developers should ensure that buffers are restricted so they can't be exploited

Prevalent vulnerabilities in 2017 included buffer overflow errors, Apache Struts Buffer overflow errors topped the list of Common Weakness Enumeration (CWE) vulnerabilities tracked by Cisco in 2017, although other categories showed movement up and down. Input validation vulnerabilities increased, while buffer errors declined (Figure 34)

GT: the same as the cybersecurity RAG one

A: Because of their prevalence, buffer overflows have been a popular target for attackers. In fact, according to the Common Weakness Enumeration (CWE), buffer overflows are the most common type of coding error exploited by criminals. To prevent this error, developers should ensure that buffers are restricted so they can't be exploited.

12. **Q:** What are the systemic issues or mistakes made by defenders that contributed to attackers' success?

These concepts might be useful to answer the question:

C: 48Recommendations For Defenders

While reviewing 2023's incidents, we kept in mind two topics close to defenders' hearts; factors that contributed to attackers' success, and mitigations defenders can use to frustrate attackers. This section distills the most critical lessons into actionable recommendations for defenders. These are pragmatic, effective approaches to counter emerging and evolving cyberthreats18Incident Response Report 2024 18 Spotlight On: Speed

Clearly, defenders need to speed up as much as possible. We'll give more recommendations for how to do that later in this report, but for the moment, there are a few key principles to keep in mind:

- **Preparation.** One of the best ways to get ahead of attackers is to truly get ahead. Through proactive preparation, organizations can get ready to respond within hours to compromises, stopping attackers before they have a chance to execute their plans.
- **Automation.** As we mentioned in the Black Basta example above, human security teams need to sleep, and they're also typically pulled in many directions. Automation—especially when powered by machine learning and AI—can help defenders sift through alerts and surface the ones that truly need attention.
- **Zero Trust.** Another powerful way to limit the damage an attacker can do after compromise is to limit their movement and activity. When organizations design their security posture in accordance with a Zero Trust philosophy, attackers become less powerful when they gain initial access, because initial access means much less. It's the difference between a thief getting into an entryway and being able to move through the entire building, and a thief getting into the lobby only to encounter another locked door.

- **Defense in depth.** A security program designed with overlapping defenses and controls gives attackers more opportunity to alert you to their presence. Especially combined with limited privileges in a Zero Trust philosophy, you can raise the signal-to-noise ratio of meaningful alerts that will let you focus on attacker activity earlier in the attack lifecycle.

GT: the same as the cybersecurity RAG one

A: The systemic issues or mistakes made by defenders that contributed to attackers' success include: 1. Lack of preparedness: Defenders often fail to prepare adequately for potential attacks, leaving them vulnerable to sophisticated attacks. 2. Slow response time: Defenders take too long to detect and respond to attacks, giving attackers ample time to carry out their malicious activities. 3. Insufficient resources: Defenders may not have enough resources to effectively monitor and protect their systems, making it easier for attackers to exploit vulnerabilities. 4. Poor communication: Defenders may not communicate effectively with each other or with external parties, leading to delays in incident response and increased risk of data breaches. 5. Inadequate training: Defenders may lack the necessary skills and knowledge to identify and respond to complex attacks, leaving them vulnerable to attackers who are better equipped. 6. Overconfidence: Defenders may underestimate the capabilities of attackers, leading to complacency and a false sense of security. To mitigate these issues, defenders should prioritize preparation, invest in automation and AI-powered tools, implement a Zero Trust architecture, adopt defense-in-depth strategies, improve communication and collaboration, and provide ongoing training and education to their teams.

13. **Q:** What are Tactics, Techniques and Procedures (TTPs)? How they are linked with the concept of 'exploit target'?

These concepts might be useful to answer the question:

TTP, ExploitTarget

C: Tactics, Techniques and Procedures (TTP) are representations of the behavior or modus operandi of cyber adversaries.

Source: https://stix.mitre.org/about/documents/STIX_Whitepaper_v1.1.pdf

ExploitTargets are vulnerabilities or weaknesses in software, systems, networks or configurations that are targeted for exploitation by the TTP of a ThreatActor.

Source: https://stix.mitre.org/about/documents/STIX_Whitepaper_v1.1.pdf

GT: the same as the cybersecurity RAG one

A: Tactics, Techniques and Procedures (TTPs) are the methods used by threat actors to achieve their objectives. They can include malware, social engineering techniques, and other means of gaining access to systems or data. Exploit targets are specific vulnerabilities or weaknesses in systems or applications that can be exploited by threat actors using TTPs. These targets may include unpatched software, misconfigured firewalls, or weak passwords. By understanding both TTPs and exploit targets, organizations can better protect themselves from cyber threats.

14. **Q:** Provide a definition of an 'attack'. What type of supply chain attack occurred in September 2017?

These concepts might be useful to answer the question:

Attack, Means, Consequence, Attacker, BotnetAttack, MeansOrConsequence, MitnickAttack

C: Researching and identifying infrastructure allows organizations to identify the following: Tactical intelligence to use in an ongoing incident response effort, infrastructure used by adversaries before it's used against the organization, and historical context or associations between infrastructure and attackers. 20 For more details on this investigation, see the RSA report, "Kingslayer—A Supply Chain Attack": [rsa.com/en-us/resources/kingslayer-a-supply-chain-attack](https://www.rsa.com/en-us/resources/kingslayer-a-supply-chain-attack). Figure 29 Links among infrastructure used by APT group Supply chain attacks: One compromised vector can affect many organizations Much like any enterprise looking to save time and money, attackers seek out ways to make their operations more efficient. As the Cisco partner RSA discovered, supply chain attacks offer maximum impact for minimal effort on the part of criminals. Another supply chain attack, which occurred in September 2017, involved the download servers used by a software vendor to distribute a legitimate software package known as CCleaner.7 CCleaner's binaries, which contained a Trojan backdoor, were signed using a valid certificate, giving users false confidence that the software they were using was secure. The actors behind this campaign were targeting major technology companies where the software was in use, either legitimately or as part of shadow IT. Supply chain attacks appear to be increasing in velocity and complexity. They can impact computers on a massive

scale, and can persist for months or even years. Defenders should be aware of the potential risk of using software or hardware from organizations that do not have a responsible security posture. Look for vendors that issue CVEs, are quick to address vulnerabilities, and consistently strive to ensure that their build systems can't be compromised. What they need from security vendors is the "how"—that is, how to implement a layered approach to value-chain security that includes elements such as physical security and cybersecurity standards. The complexity of their networks means utility and energy organizations must also assess the impact of threat alerts, and decide which ones deserve mitigation resources. Nearly half of the utility and energy security professionals said of the thousands of daily alerts they see, only 63 percent of those alerts are investigated. Of the alerts that are investigated, 41 percent are deemed legitimate threats, and 63 percent of those threats are remediated. Figure 68 Targeted attacks and APTs are most critical concerns

Download the 2017 graphics at: cisco.com/go/mcr2017graphics Source: Cisco 2017 Security Capabilities Benchmark Study 42% 40% Targeted Attacks Advanced Persistent Threats Utilities Figure 68 Targeted attacks and APTs are most critical concerns

GT: the same as the cybersecurity RAG one

A: An attack is a hostile action against a person or thing. In September 2017, a supply chain attack occurred in which a software vendor distributed a legitimate software package known as CCleaner, but the binaries contained a Trojan backdoor that was signed with a valid certificate, giving users false confidence that the software they were using was secure. The actors behind this campaign were targeting major technology companies where the software was in use, either legitimately or as part of shadow IT.