

# 次世代学術情報基盤に向けた基盤ソフトウェアの 実践的な研究・開発・評価

空閑 洋平（東京大学）

## 概要

本プロジェクトは、2021年度から本格稼働したNII Research Data Cloud (RDC) と高性能仮想化基盤 mdx、情報基盤センターや HPCI の大型計算機群との連携を深化させ、次世代の計算科学・データ科学・データ駆動科学・データ利活用分野を支える学術情報基盤に向けた基盤ソフトウェアの研究、開発、評価を行う。本課題は昨年度からの継続課題であり、本年度は、昨年度設定した研究テーマを継続実施すると共に、次の段階として (1) mdx の機能強化として、クラウド間データ転送ソフトウェアの開発、mdx における脆弱性試験の導入、VM テンプレート機能の開発、(2) NII RDC と mdx 間の連携強化として、mdx 上の JupyterHub 環境と GakuNin RDM の連携システムの構築支援機能の開発、OpenOnDemand を含む HPC クラスターの構築支援機能の開発を実施した。それぞれの成果は、論文、オープンソースの公開、サービスの稼働を実施した。

## 1 共同研究に関する情報

### 1.1 共同研究を実施した拠点名

- 北海道大学 情報基盤センター
- 東京大学 情報基盤センター
- mdx

### 1.2 課題分野

- データ科学・データ利活用課題分野

### 1.3 共同研究分野 (HPCI 資源利用課題のみ)

- 超大規模情報システム関連研究分野

### 1.4 参加研究者の役割分担

- 空閑 洋平（東京大学、研究代表者）：仮想化基盤・ネットワークに関する研究開発
- 中村 遼（東京大学、副代表者）：ファイル転送ツールの開発
- 杉木 章義（北海道大学）：遠隔転送の拠点

### 間調整

- 宮本 大輔（東京大学）：セキュアデータ環境の実現
- 埴 敏博（東京大学）：mdx 計算機との調整 (GPU 連携機能、スパコンとの接続)
- 田浦 健次朗（東京大学）：mdx 計算機との調整
- 合田 憲人 (NII)：RDC 等連携機能の開発
- 竹房 あつ子 (NII)：RDC 等連携機能の開発
- 藤原 一毅 (NII)：RDC 等連携機能の開発
- 大江 和一 (NII)：RDC 等連携機能の開発

## 2 研究の目的と意義

2021年度から本格稼働したNII Research Data Cloud (RDC) と高性能仮想化基盤 mdx、情報基盤センターや HPCI の大型計算機群と

の連携を深化させ、次世代の計算科学・データ科学・データ駆動科学・データ利活用分野を支える学術情報基盤に向けた基盤ソフトウェアの研究、開発、評価を行う。本課題は、昨年度からの継続課題であり、本年度は、昨年度設定した研究テーマを継続実施すると共に、次の段階としてデータ連携機能のソフトウェア化・サービスの稼働を目指している。

### 3 当拠点公募型研究として実施した意義

現在各拠点では、データ科学分野の研究を推進するために拠点間、拠点・クラウド間のデータ連携機能の研究・開発を実施しており、このような拠点間の相互接続性強化のためには、各拠点の研究者で連携する体制が重要となる。本課題は、募集課題分野2「データ科学・データ利活用課題分野」の研究を推進するための学術基盤の機能強化を目的としたネットワーク型の拠点連携課題であり、各拠点の資源利用ができる当拠点公募型共同研究で実施することが適していると考えている。各研究者が個別に研究してきた成果を持ち寄り、mdx を中核とした実環境にデプロイすることで、研究環境の整備を進めるとともに、現状の課題や問題点を明らかにする。

### 4 前年度までに得られた研究成果の概要

昨年度公開した論文とオープンソースの概要と、今年度の進展について述べる。

昨年度の mdx でのデータ蓄積・解析基盤の事前検討で実施した東大 Zoom データの保存、解析環境の構築に関する論文は、今年度に継続成果を論文誌 1 編に投稿し、また昨年度の研究会論文が論文賞を 2 件受賞した [1-1] [6-1] [6-2]。また、昨年度オープンソースを公開し

たクラウド間データ転送ソフトウェア mscp [5-1] に関しては、現在も開発を継続し、今年度は、査読あり国際会議 1 編、査読なし国内会議 1 編、Data Mover Challenge 2023 に参加し、Most Innovative for HPC users Award を受賞した [2-1] [4-1] [6-3]。また、昨年度公開した、NII RDC と mdx の連携強化に関するオープンソースプロジェクトについては、こちらも開発を継続し、mdx と GakuNin RDM のシームレスな連携機能、Open OnDemand 環境の構築方法を検討し、成果を査読なし国内会議 3 編、オープンソースで 3 件公開した [4-4] [4-2] [4-3] [5-2] [5-3] [5-4]。

また、今年度の新規取り組みとして目指していたサービスの稼働関係については、mdx 上の脆弱性診断機能の検討、VM テンプレート機能追加についてサービス化を実施したため、次節で詳細を述べる。

## 5 今年度の研究成果の詳細

### 5.1 テーマ 1: 仮想環境間のデータ連携強化

#### 5.1.1 仮想環境間的高速かつ手軽なデータ転送

本テーマでは、手軽に高速大容量のデータ転送を実現するツールとして、multi-threaded scp (mscp) を開発した。mscp は、mdx のように、ユーザが自身の手で仮想マシンにソフトウェアをインストール、設定、チューニングしなければならない環境においても、簡単に利用できることを主眼に設計されている。

mscp は複数の SSH コネクションごとにファイルを分割して転送することで、sshd という広く利用されているサーバソフトウェアに対して高速にデータを転送する。2023 年度の前半は、mscp の性能について評価を行い、図 1 に示すように、100Gbps で直接接続された 2 台のマシン間でのファイル転送で 44Gbps のスループッ

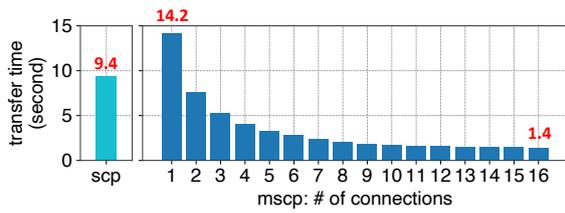


図1 100Gbps で接続された2台のコンピュータ間での mscp と scp のスループット計測結果 ([2-1] より)

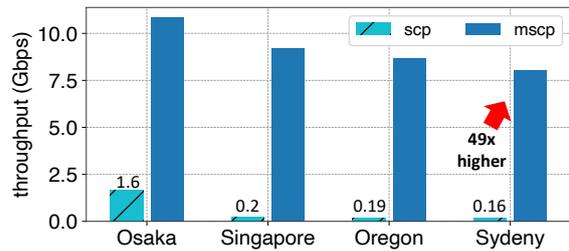


図2 Google Cloud 上での東京から各リージョンへのファイル転送におけるピークスループットの比較

トを達成した。また Google Cloud を用いた大陸間のファイル転送では、東京からシドニーへのファイル転送で scp コマンドの 49.3 倍のピークスループットを達成した。図2にこの結果を含む他のリージョンへのピークスループットを示す。これらの成果は ACM の Practice and Experience in Advanced Research Computing (PEARC) に査読付き論文として採録され、発表を行った [2-1]。また Supercomputing Asia の併催イベントである Data Mover Challenge 2023 (<https://www.nsc.sg/data-mover-challenge-2023/>) に参加し、世界 11 箇所に設置され 100Gbps で接続されたノード間でのデータ転送を行った。本転送実験において mscp は scp と比較して、オーストラリアから日本へのデータ転送では、約 240 倍、サウジアラビアからフィンランドへの転送では約 131 倍高速にファイルを転送することができた。以上の結果をもって、Most Innovative for HPC uses Award を受賞した [6-3]。

mscp のソースコードは Open Source で公開している [5-1]。また主要な Linux Distribution 向けにはバイナリパッケージも公開しており、誰でも広く利用可能である。2023 年夏の v0.1.0 公開以降、すでにいくつかの組織や個人から利用の報告や機能の追加要望を受けており、今後も開発とメンテナンスを継続してい

く予定である。

### 5.1.2 mdx VM テンプレート機能の強化

mdx のような IaaS 基盤は、利用者が独自にクラスタを構築可能な一方で、Linux の運用に不慣れた利用者にとっては、複数仮想マシンで構成されるクラスタの構築の難易度は高く、クラスタの構築、運用支援が課題となっている。そこで、本年度は、クラスタ構築と、特定研究分野に特化した、設定済みの OS イメージである仮想マシンテンプレート提供の開発を実施した。具体的には、特定研究分野向けの仮想マシンテンプレートとして MateriApps LIVE! イメージの仮想マシンテンプレート化した\*1。また、mdx 上での NFS や LDAP を含む複数仮想マシンによるクラスタ構築を目的とした仮想マシンテンプレートを開発し、利用者向けに公開した [図 3]。

MateriApps VM では、東京大学物性研究所計算物質科学研究センターが公開している、計算物質科学アプリケーションの利用環境を、mdx の VM テンプレートとしてパッケージした VM テンプレートである。MateriApps LIVE!環境は、手元 PC 環境などでの利用を想定したもので、MateriApps VM テンプレート

\*1 <https://ma.issp.u-tokyo.ac.jp/materiapps-live>

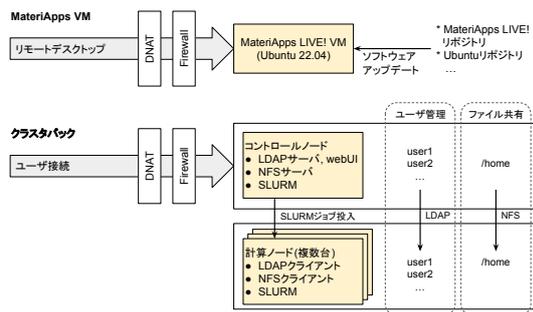


図3 上図: MateriApps VM テンプレートの概要. 下図: クラスタパッケの概要

を用いることで、同様の環境 mdx 上に構築できる。利用者は、手元 PC からリモートデスクトップ機能を用いて mdx の VM に接続可能になっている。

クラスタパックは、mdx 上に複数マシンからなる CPU, GPU クラスタの構築、管理を支援する VM テンプレートである。クラスタパックは、コントロールノード VM と、計算ノード VM テンプレートの2つで構成される。コントロールノード VM は、クラスタ内に1台のみデプロイし、利用者のログインノード、ユーザアカウント管理のための LDAP サーバ、共有ディレクトリのための NFS サーバ、SLURM によるジョブシステムがパッケージされている。また、利用者は、LDAP Account Manager を用いて、Web ブラウザによる GUI でユーザアカウントの管理が可能になっている。計算ノード VM は、コントロールノードを参照するクライアントとして動作する。クラスタ内に、必要な量の CPU, GPU ノードを計算ノード VM テンプレートを用いてデプロイすることで、クラスタが構築できる。デプロイされた計算ノードは、SLURM のインストール、LDAP ユーザ、NFS による /home が共有されており、コントロールノードから SLURM によるジョブの投入や ssh ログインが可能になっている。

## 5.2 テーマ 2: セキュアデータ環境, 仮想化環境の高機能化の実現

本テーマでは、mdx におけるセキュリティ機能の開発と実装方法を検討しており、本年度は mdx における CSERT チームの発足と、ユーザ環境への脆弱性検査を実施した。

### 5.2.1 mdx-CSIRT の発足

mdx におけるインシデントは、システムを提供する「資源提供者」側、システムを利用する「利用者」側の両方で責任を分担する、いわゆる責任分担モデルを採用している。例えば仮想化ソフトウェアやポータルに脆弱性がある場合は資源提供者側の問題であり、資源提供者が主体となってインシデント対応を行う。一方で、VM 内に脆弱性のあるソフトウェアを稼働させている場合は利用者の問題となる。いずれの場合も、インシデントの早期発見と早期対応が必要であることは重要である。

mdx-CSIRT は、mdx におけるインシデントの対応を行う専門的なチームである。実際の業務としてはサービステスト時よりすでに稼働していたが、正式に mdx-CSIRT の内規を制定することで、CSIRT を発足させた。

本 CSIRT の内規の特徴は以下のとおりである。

**統括責任者・メンバーの選出** CSIRT の統括責任者を内規に沿って選出されるようにした。また、CSIRT メンバーをデータ活用社会創成プラットフォーム協働事業体の構成機関に所属する者を中心として統括責任者によって選出されるようにした。インシデント発生時には、統括責任者より本インシデントの担当者を割り当て、担当者がインシデントの原因、影響範囲を調査し、問題の対応に当たる。原因が利用者側にある時は、適切な窓口となる CSIRT メン

バーを選出する。インシデント時のコミュニケーションは重要であり、特に問い合わせを行う者と受ける者との間の信頼関係の醸成は必須と言える。利用者がデータ活用社会創成プラットフォーム協働事業体の構成機関との共同研究により mdx を利用している場合には構成機関に所属するメンバーから連絡をする等、コミュニケーションには配慮が必要と考えられる。

**情報取扱規定** 情報取扱規定について、構成機関が機密性ラベルを採用している一方で、格付けの区分、格付けの対象範囲、取扱制限などに違いも見られる。そこで、一般的な CSIRT が採用している情報取り扱い規定である Traffic Light Protocol (TLP) を採用し、TLP:Red ~ TLP:Clear までの情報取扱規定を設定した。

この他、インシデント発生時の緊急的な措置を想定した統括責任者の権限等が記載されている。

CSIRT 発足後の 2024 年 3 月には第 1 回目となるミーティングを開催し、2024 年度にも 3 回以上のミーティングを予定している。また、今後は利用者向けに mdx のセキュアな利用についての啓蒙活動を行う予定である。

### 5.2.2 mdx 仮想化環境の脆弱性診断機能の実現

mdx では、利用者が仮想マシンを作成し占有利用可能になっているため、脆弱性がある状態で VM が運用される可能性がある。作成した仮想マシンのネットワークは、基本的に NAT 環境下に配置されるため外部からアクセスできないが、DNAT 機能を使うことで、仮想マシンにグローバル IP アドレスを付与できる。そこで、本テーマでは、グローバル IP アドレスを利用している全ての仮想マシンに対して脆弱性

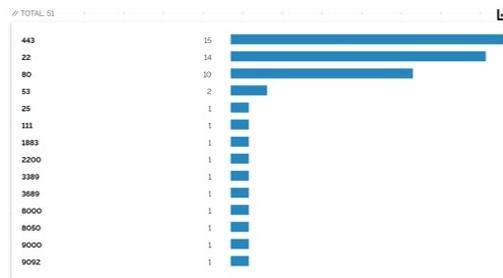


図 4 mdx に対するポート番号別スキャン行為の有無

検査を実施する目的でサービスを構築した。

はじめに、mdx のグローバル IP アドレスレンジへのスキャン行為の有無の調査を実施した。図 4 は 外部サイト SHODAN(\*2) を用いて調査した、mdx へのスキャン行為の有無を調査した結果を表す。本調査により、mdx に対する第三者のスキャン行為は継続的に行われていることを確認し、定期的なセキュリティ検査の必要性を確認した。また、mdx では、すでにプロジェクト利用者の mdx 上での計算資源を悪意の第三者がサイバー攻撃により不正に利用する事例が報告されている。

次に、今年度実施したセキュリティ検査の結果を報告する。脆弱性検査には Greenbone Vulnerability Management(GVM) を用い、既知の脆弱性情報に基づいた診断を行った。当初は希望するプロジェクトに対してのみ実施することを検討していたが、先述のインシデント情勢を受けて全てのプロジェクトに対して実施し、希望しないプロジェクトについては連絡により受け入れる方針とした。なお、脆弱性診断の通知日は 2023 年 7 月 28 日、不参加の希望の連絡期日は 2023 年 8 月 31 日、診断の実施日は 2023 年 9 月 11 日～15 日とし、脆弱性診断を行うホストの IPv4 アドレスとともに通知

\*2 <https://www.shodan.io/>

	アドレス数	プロジェクト数
総数	225	105
検査対象外	46	23
検査対象	179	82

表 1 脆弱性検査の対象について

した。

脆弱性検査の対象は DNAT 用に払い出されたグローバル IPv4 アドレスである。脆弱性検査対象ポートは ACL から判別可能であればこちらを、判別が難しい場合は GVM における標準的な検査設定である「全ての TCP ポート及び著名な UDP ポート」を用いることとした。なお、mdx の現行の環境では、明示的に ICMP プロトコルに応答するような設定を行わないと ICMP ECHO Request に対する応答は行われない。この場合、GVM はホストが実際に稼働しているにも関わらず、検査対象外とみなす仕様がある。この仕様を回避するための設定等、様々な最適化が必要であると考えられた。

そこで、DNAT 及び ACL のリストから、自動的に GVM を設定する方針を採用した。GVM には専用の API 及びコマンドライン操作ツールが用意されている。DNAT 及び ACL の設定から、各プロジェクト毎に「アセット」として検査対象 IP アドレスおよび検査予定ポートを登録し、自動的に最適化された検査を登録するプログラムを作成し、これによって自動的に GVM の設定を行った。

脆弱性検査の対象は表 1 に示す通り、82 プロジェクトが利用する 179 個の IPv4 アドレスに対して行われた。検査の実施は自動ではなく、サーバの負荷及び NAT を行うネットワーク機器の負荷の影響を見ながら手動によって行った。なお検査機器は mdx における CPU パックを使用し、不要な部分は返却している。

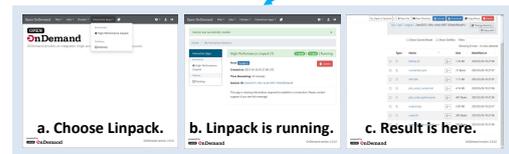
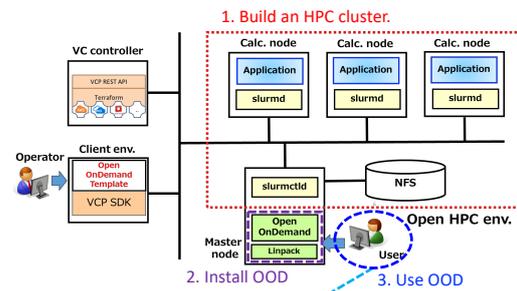


図 5 mdx 上に構築した Open OnDemand 環境の概要。

GVM は、発見された脆弱性の影響を High, Medium, Low の 3 種類で分類する。High に分類されるものは緊急対応が必要であると考えられる脆弱性であるが、今回は発見されなかった。Medium に分類された対応が必要となる可能性があると思わしき脆弱性は 3 件発見され、各プロジェクトの代表者に通知により連絡した。

今後は必要に応じて対応の監査を行い、また得られた知見の共有を行っていく。近年は AI の研究者が 38 テラバイトのデータを漏洩する\*3事案が発生しており、クラウドの不正な設定およびインシデントには枚挙に暇がない。2024 年度は脆弱性診断サービスを拡充し、年に複数回実施する予定である。

### 5.3 テーマ 3: NII RDC, 他システムとのデータ連携・情報共有

#### 5.3.1 RDM 連携, Open OnDemand 対応, および活用事例の検討

本テーマでは, 昨年度から継続して, NII RDC と mdx 間, および mdx と基盤センタースパコンとの連携, およびそれらの活用事例の検討を進めている. NII では, クラウド資源上でコンテナと Jupyter Notebook を活用して研究・教育アプリケーション環境を容易に再現する VCP システムの開発を進めており, 昨年度は mdx API と VCP で講義演習環境の構築を実証した. 今年度は, mdx で VCP を用いてユーザプロジェクトごとの (1)NII RDC 連携, (2) スパコン活用支援, および (3) 活用事例の検討を進めている.

(1)NII RDC 連携では, VCP を用いて mdx 上で構築した JupyterHub 環境 (The Littlest JupyterHub) に対して, RDC の研究データ管理基盤である GakuNin RDM の解析機能としてシームレスに利用できるようにし, その動作を確認した.

(2) スパコン活用支援では, 図 5 に示すように VCP を用いて mdx 上でオートスケーリング可能な HPC クラスタを構築するとともに, スパコンのポータルとして利用されている Open OnDemand 環境を配備できるようにした [4-2][4-3]. また, 新たに OpenHPC ライブラリ v3 に対応した HPC クラスタの構築も行えるようにした. これにより, Open OnDemand の Web UI 経由で mdx 上に構築したクラスタ環境を簡単な操作で利用できるようにした. また, mdx 上に構築した HPC 環境と, ABCI などの他のスパコンをシームレスに利用するため

の連携動作方法の検討を進めた.

(3) 活用事例としては, (2) の構築手順を Jupyter Notebook ベースの手順書として GitHub サイトに公開した [5-2][5-3][5-4]. また, IoT システム構築支援ライブラリ SINET-Stream (<http://sinetstream.net/>) でセンサーデータを収集, 蓄積し, GakuNin RDM の解析機能で活用, 再利用する方法について検討し, プロトタイプ実装を進めた [4-4].

## 6 今年度の進捗状況と今後の展望

今年度は, 昨年度課題の研究開発を継続するとともに, 成果の論文化, オープンソース化, サービス化を実施し, 進捗状況としては順調な数の成果をまとめた. 今年度の論文実績としては, 論文誌 1 編, 査読あり国際会議 1 編, 査読なし国内会議 4 編を執筆した. オープンソース化実績としては, mscp, NII RDC 連携ソフトウェアで 4 プロジェクト公開した. サービス化の実績としては, mdx の脆弱性試験サービス, VM テンプレート (MateriApps, クラスタバック) の運用を開始した. また, その他受賞関係の成果としては, mscp に関しては Data Mover Challenge 2023 に参加し 1 件受賞, また, 昨年度の論文成果で情報処理学会から 2 件受賞した.

本プロジェクトは, 今後も研究開発を継続し, 論文, オープンソース, サービスの公開を継続したいと考えている.

## 7 研究業績一覧 (発表予定も含む)

### 学術論文 (査読あり)

- [1-1] 空閑 洋平, 中村 遼, “遠隔会議システムの計測データを用いたネットワーク品質計測”, 情報処理学会論文誌, 2024 年 3 月.

\*3 <https://www.wiz.io/blog/38-terabytes-of-private-data-accidentally-exposed-by-microsoft-ai-researchers>

#### 国際会議プロシーディングス (査読あり)

- [2-1] Ryo Nakamura and Yohei Kuga. 2023. “Multi-threaded scp: Easy and Fast File Transfer over SSH”. In Practice and Experience in Advanced Research Computing (PEARC '23). ACM, New York, NY, USA, 320–323. <https://doi.org/10.1145/3569951.3597582>

#### 国際会議発表 (査読なし)

#### 国内会議発表 (査読なし)

- [4-1] 中村 遼, 空閑 洋平, “SSH による高速ファイル転送ライブラリの開発”, 電子情報通信学会, インターネットアーキテクチャ研究会, 10月研究会 (ADVNET 共催), 2023.
- [4-2] 大江 和一, 竹房 あつ子, 丹生 智也, 合田 憲人, “クラウド環境構築システム VCP による mdx でのスケーラブルな HPC クラスターの構築”, 情報処理学会研究報告 2023-HPC-190, No. 9, pp. 1-7, 2023.
- [4-3] 大江 和一, “クラウド環境構築支援システム VCP による Open OnDemand 環境の簡易構築”, PCCC Open OnDemand ワークショップ, 2023.
- [4-4] 藤原 一毅, 竹房 あつ子, 北川 直哉, 小林 久美子, 合田 憲人 “IoT 実験のための SINETStream お試しサービスの構想”, インターネットと運用技術シンポジウム (IOTS2023) (ポスター), 2023.

#### 公開したライブラリ等

- [5-1] mscp: multi-threaded scp, <https://github.com/upa/mscp>.
- [5-2] 学認クラウドオンデマンド構築サービス 計算資源補完のためのテンプレート, <https://github.com/nii-gakunin-cloud/ocs-templates/tree/master/HybridCloud>.

- [5-3] 学認クラウドオンデマンド構築サービス Open OnDemand テンプレート, <https://github.com/nii-gakunin-cloud/ocs-templates/tree/master/OpenOnDemand>.

- [5-4] 学認クラウドオンデマンド構築サービス OpenHPC v3 テンプレート, <https://github.com/nii-gakunin-cloud/ocs-templates/tree/master/OpenHPC-v3>.

#### その他 (特許, プレス発表, 著書等)

- [6-1] 空閑 洋平, 中村 遼, 情報処理学会 (IOT 研究会) 藤村記念ベストプラクティス賞, 2023年7月.
- [6-2] 空閑 洋平, 情報処理学会 山下記念研究賞, 2024年3月.
- [6-3] Ryo Nakamura, Yohei Kuga and Kunio Akashi, Data Mover Challenge 2023 Most Innovative for HPC uses Award, [https://www.u-tokyo.ac.jp/focus/ja/articles/z0310\\_00050.html](https://www.u-tokyo.ac.jp/focus/ja/articles/z0310_00050.html)