

RPAの導入と使い方

2020年6月22日

Quark

はじめに

- 複数のPDF、オフィス文書、WEBページ等を対象に、全ページを連続コピーまたは指定部分をOCR（文字認識）で連続抽出し、Excelリスト化する機能です
 - Copy Full Text
 - ✓ 複数文書のExcelリスト化（全ページ連続コピー PDF、Word、他）
 - OCR Selected Region
 - ✓ 複数文書のExcelリスト化（指定エリア連続抽出 PDF、Word、画像、他）
- 必要システム
 - RPAをお使いの場合には、64bitのWindows環境が必要です
 - ユーザフォルダ名が日本語の場合、不具合の恐れがあります
 - ✓ マイクロソフトアカウント名が日本語になっている、もしくは、今は英字アカウントだが、初期設定で日本語アカウントを使っていた場合、動作しない可能性があります
 - macには対応していませんが、Windowsの仮想環境での動作は確認しています
- ご注意
 - RPA実行中は、原則として他のPC作業はできません
 - OCRの精度を上げるため、ご使用ソフトの設定は、全画面かつページ幅に合わせることをおすすめします
 - OCRは時間を要しますので、指定する範囲、ファイル数にご注意下さい
 - OCRで起点画像が見つからない場合、1分間探し続けますがフリーズではありません

環境設定

■ Javaの実行エンジンのインストール

- 必ず64ビット版をインストールして下さい
- <https://java.com/ja/download/manual.jsp>

ど、特定の使用が無償で許可されますが、以前のOracle Javaライセンスの下で認可されたその他の使用はできなくなりました。この製品をダウンロードして使用する前に、条件をよくお読みください。FAQがこちらでご覧いただけます。

商用ライセンスおよびサポートは、安価なJava SEサブスクリプションにより提供されます。

オラクル社は、オープン・ソースのGPLライセンスの下、jdk.java.netでOpenJDKの最新リリースも提供しています。

ご使用のコンピュータ用に最新のJavaを入手するには、オペレーティング・システムに適したファイルを下のリストから選択してください。

> [旧バージョンの削除](#)

> [Javaとは](#)

Javaのダウンロードにより、[Oracle Java SEのOracle Technology Network使用許諾契約の条項](#)を読み、同意したことを承認することになります

64ビット版

 Windows  [どれを選択すればよいですか。](#)

 Windows オンライン ファイルサイズ: 1.97 MB	手順	Javaをインストールしてから、ブラウザでJavaを有効にするには、ブラウザを再起動する必要があります。
 Windows オフライン ファイルサイズ: 65.52 MB	手順	
 Windows オフライン (64ビット) ファイルサイズ: 73.73 MB	手順	

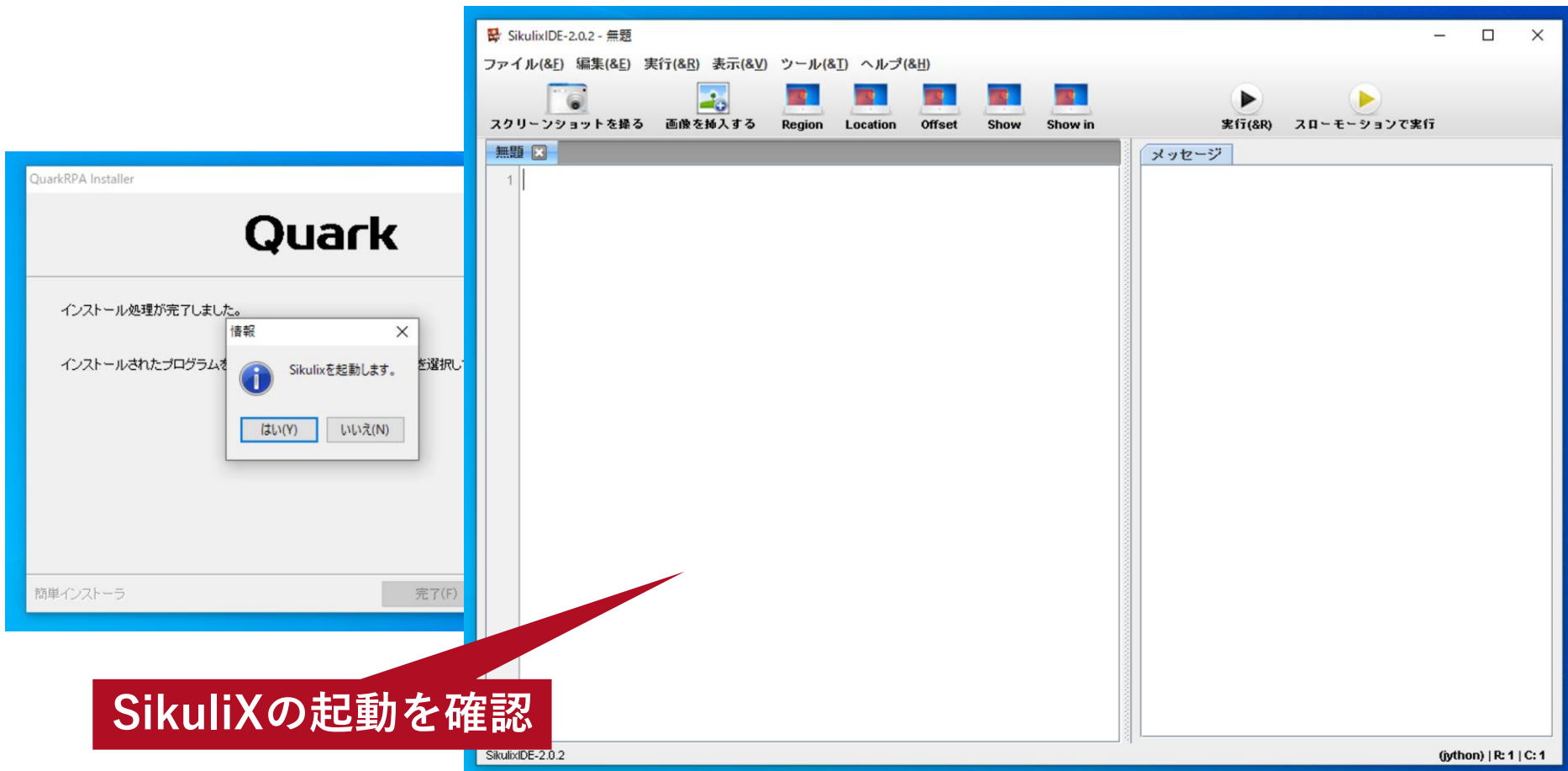
32ビットと64ビットのブラウザを切り替えて使用している場合、両方のブラウザでJava Plug-inを使用できるように、32ビットと64ビットの両方のJavaをインストールする必要があります。 » [Windows版64ビットJavaに関するFAQ](#)

 Mac OS X  [MacのFAQ](#)

環境設定 つづき

■ RPA関連ツールのインストール

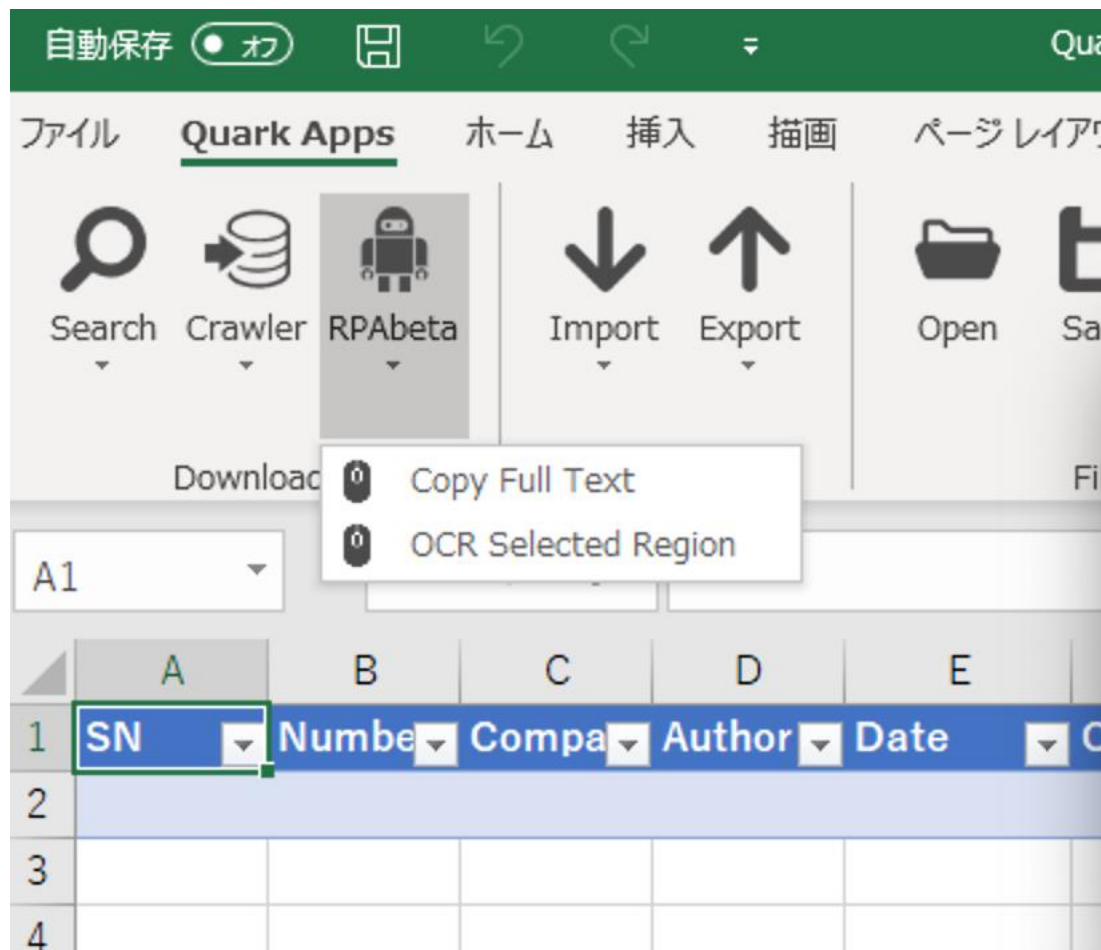
- 弊社サポートサイトから、SikuliX、日本語学習データ等をパッケージ化したインストーラーをダウンロードして下さい
- <https://support.qrk.co.jp/downloads>
- インストール後、指示通りにSikuliXを起動して下さい、SikuliXが起動したら完了です



SikuliXの起動を確認

RPAの起動 ・ ・ Copy/OCR共通

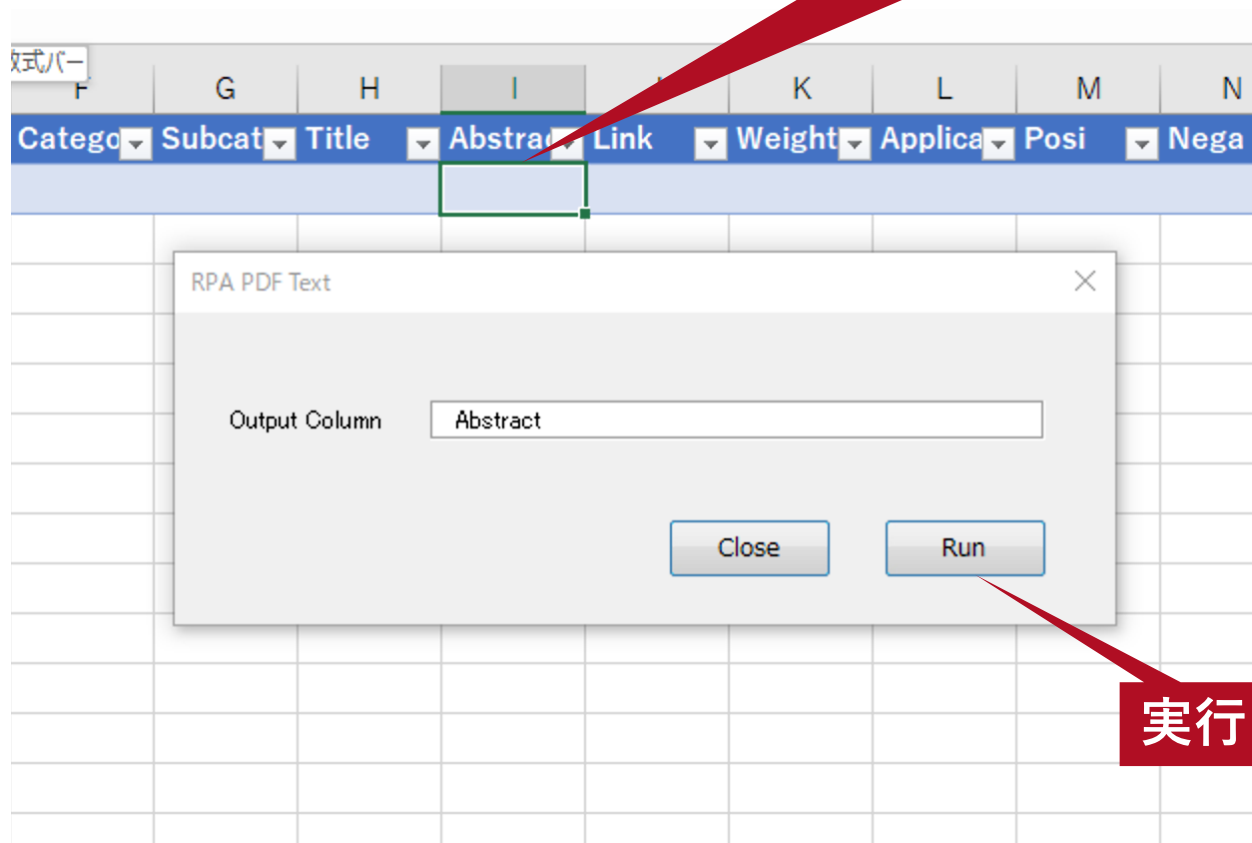
- Quark AppsのRPAbetaボタンから、Copy Full TextかOCR Selected Regionを選択



出力先の指定 ・ ・ Copy/OCR共通

- 出力先の列をクリックしてRun

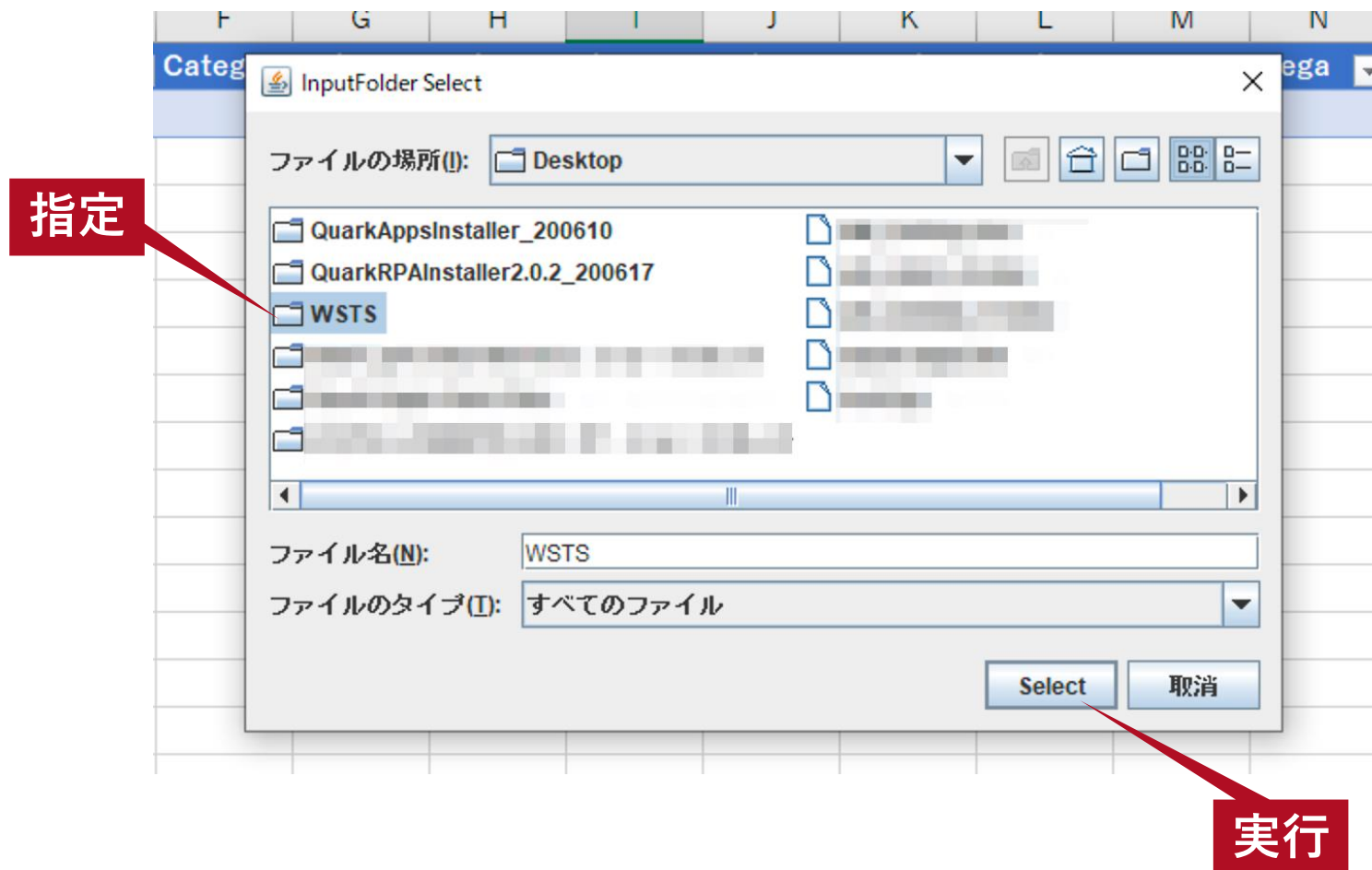
出力先(列)をクリック



実行

ドキュメントフォルダの指定 ・ ・ Copy/OCR共通

- リスト化したいドキュメントが格納されているフォルダを指定します



- Copy Full Textの操作はここまで、連続コピーがスタートします

起点となる画像の指定 ・ ・ OCRのみ

- フォルダ内の最初のファイルが開いて待機状態となります
- 起点となる画像のあるページまでスクロールし、OKをクリックします

ものの、年間では同+15.9%と2年連続の1桁成長を予測した。

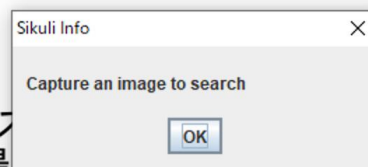
2019年は米中貿易摩擦を筆頭に世界経済に不透明要素が多いことから慎重な見方となり、半導体市場合計で前年比+2.6%と予測した。

なお、US\$1に対する円の為替レートは、2017年：112.1円、2018年：110.1円、2019年以降は111.5円を前提としている。

○ 製品別市場動向(世界市場)

2018年における製品別のドルベースでは、ディスプレイは前年比+11.7%、市場規模242億ドル、オプトは同+11.2%、市場規模134億ドル、IC全体は同+17.0%、市場規模4,016億ドルと予測した。ICの製品別では、メモリは前年比+33.2%、ロジックは同+7.3%、マイクロは同+6.4%、アナログは同+10.8%と予測した。

2019年には、ディスプレイは前年比+3.9%、市場規模251億ドル、オプトは同+6.8%、市場規模414億ドル、センサーは同+5.1%、市場規模141億ドル、IC全体は同+2.0%、市場規模4,096億ドルと予測した。ICの製品別予測では、メモリは前年比-0.3%、ロジックは同+3.8%、マイクロは同+3.0%、アナログは同+3.8%と予測した。



○ 日本の半導体市場動向

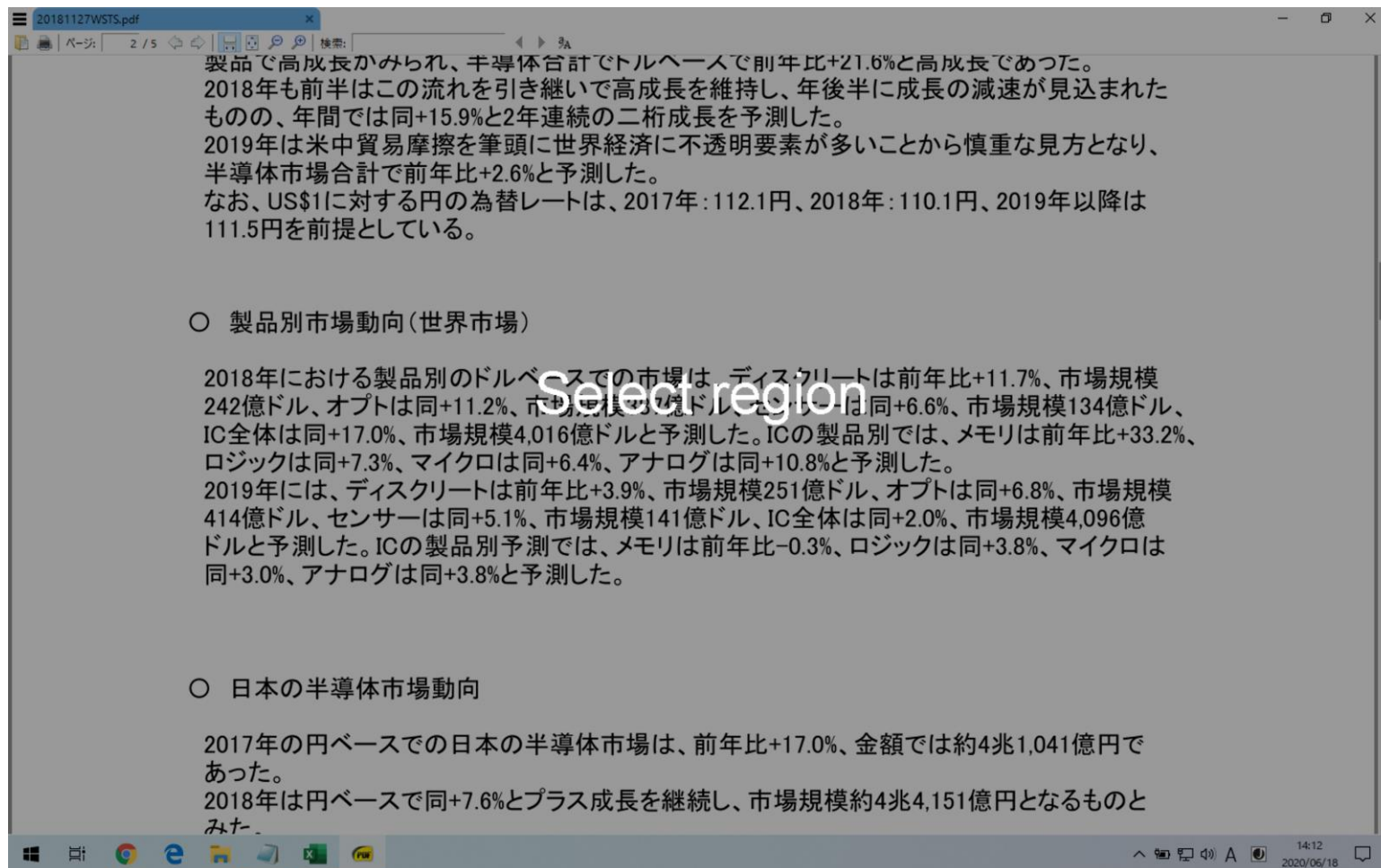
起点となる画像の指定 ・ ・ OCRのみ つづき

- “capture an image”が表示されたら、起点となる画像領域を指定して下さい



抽出する領域の指定 ・ ・ OCRのみ

- “Select region”が表示されたら、抽出したい領域を指定して下さい
 - 起点画像からの相対位置をシステムが記憶します



- OCR Selected Regionの連続抽出がスタートします

結果

- 処理が終わると、指定した列に情報が入力されます
 - Excelの仕様により、セル内の最大文字数は約3万文字、ワード文書だと約20ページです

Download

Data

File

Text

Visualization

I2

✕

✓

$\mathcal{f}_x$

2018年における製品別のドルベースでの市場は、ディスクリットは前年比242億ドル、オプトは同+11.2%、市場規模387億ドル、センサーは同+6.6%、

</

Quark

クォーク株式会社

141-0001 東京都品川区北品川5-5-15 大崎ブライトコア4F

T 03-5755-5015 F 03-3722-3410

<https://qrk.co.jp>